

# **NOVEL FUNCTIONS OF THE MOZ DOUBLE PHD FINGER DOMAIN**

**Sian Elizabeth Deeves, MBiochem (Hons).**

Thesis submitted to the University of Nottingham for the  
degree of Doctor of Philosophy

**September 2011**

## Abstract

Monocytic leukaemia zinc-finger protein (MOZ) is a histone acetyltransferase (HAT) implicated in haematopoiesis and acute myeloid leukaemia, as well as embryonic and postnatal development. MOZ contains multiple domains, including a MYST HAT domain and a double PHD finger domain (DPF) suggesting it interacts with histones. This work has established for the first time that the MOZ DPF exhibits dual functionality in establishing and sensing post-translational modifications (PTMs) of histones. Firstly, our data detected the direct interaction of MOZ with the N-terminal tails of histones H3 and H4 and shows that the MOZ DPF domain mediates such binding. Both PHD fingers are required and functionally cooperate to establish the DPF histone binding preference in terms of PTMs. We demonstrate that H3K4me3 prevents MOZ DPF association with H3, although H3K4me2 is tolerated. Similarly, H4Kac acts as a dominant exit signal that excludes MOZ from chromatin. This ability to sense H3K4 PTM status was confirmed in a collaborative effort establishing the crystal structure of MOZ DPF in complex with an unmodified H3 peptide. The H3 peptide adopted an  $\alpha$ -helical conformation in the complex, which has not previously been observed.

Secondly, we present novel data showing that the MOZ DPF domain exhibits a mild histone H3-specific acetyltransferase activity. This provides the first report of a possible enzymatic role in chromatin modification attributed to a PHD finger. Furthermore, the combined DPF and MYST domains were found to influence the reaction rate and substrate specificity of MOZ-induced histone acetylation.

Our studies revealed that the MOZ DPF could associate with heterochromatic PTMs, namely H3K9me3. We report here that both the H3K9-specific methyltransferase SUV39H1 and heterochromatin protein 1 (HP1) form interactions with MOZ, implicating its function in both corepressor and coactivator complexes. Thus, our data suggest that like several other chromatin-associated proteins, MOZ is a multi-functional regulator of chromatin modification and gene expression.

## Acknowledgements

I would first and foremost like to thank my supervisor, David Heery, for the opportunity to work on such an exciting project and for his continuous guidance and support over the past four years. I would also like to thank ‘my postdoc’ Hilary Collins (Hils) for investing copious amounts of time in my training, but also for her advice and friendship throughout my PhD. I am also most grateful to a number of members of the Gene Regulation Group (past and present) for all their help, friendship and for creating an enjoyable work environment. In particular, I would like to thank Marie Miller-Messmer and Rachel Doidge (whose weddings were of a great distraction), Karin Kindle, Fran Wadelin, Alex Garvin, Chun Ming Chan, Joel Fulton, Akhmed Aslam, Baigong Yue, Barbara Rampersad and Jane Doughty. I would also like to acknowledge our collaborator Ingrid Dreveny (Structural Biology) for her time and expertise in solving the crystal structure of our protein domain, Ludo Dekker (Medicinal Chemistry) for helping to set up the fluorescence polarisation protocol and Bas Winkler for his help and advice, especially with the use of structural programs. I must also thank the BBSRC for funding my PhD project.

A special thank you to my mum, dad and brother for their love and support and their encouragement through everything I have accomplished. To Oli whose love and ‘jokes’ saw me through the good and bad times. Finally, I would also like to acknowledge my nanny, of whom I miss very much, but know is very proud of what I have achieved.

# Table of Contents

<b>Chapter One: Introduction .....</b>	<b>1</b>
1.1. Transcription .....	2
1.2. Chromatin Structure and Dynamics Govern Cellular Processes .....	10
1.3. Histone Code Hypothesis .....	16
1.4. Covalent Modification of Histones and Chromatin Remodeling .....	17
1.4.1. Acetylation .....	19
1.4.2. Phosphorylation .....	20
1.4.3. Methylation .....	21
1.4.4. Ubiquitination .....	24
1.4.5. SUMOylation .....	25
1.4.6. ADP Ribosylation .....	26
1.4.7. Biotinylation.....	27
1.4.8. Proline Isomerisation .....	28
1.4.9. Histone Tail Clipping.....	29
1.4.10. Chromatin Remodelling .....	29
1.5. Lysine Acetyltransferase Enzymes .....	32
1.5.1. MYST family .....	34
1.5.2. CBP/p300 family.....	36
1.5.3. GNAT family .....	40
1.6. Structure and Function of MOZ / MYST3 .....	42
1.6.1. MOZ is a modular protein.....	42
1.6.2. MOZ resides in multi-subunit complexes .....	46
1.6.3. MOZ functions as a transcriptional activator in hematopoietic processes .....	47
1.6.4. MOZ functions as a transcriptional activator in developmental processes .....	50
1.6.5. MOZ functions as a transcriptional activator in the DNA damage response .....	52
1.7. The Role of MOZ & Other KATs in Disease .....	53
1.8. Epigenetic Modulation: Potential Anticancer Therapy .....	59
1.9. Readers of the Histone Code .....	65
1.9.1. Recognition of acetyl-lysine modifications.....	68
1.9.2. Recognition of methyl-lysine/arginine modifications .....	71
1.9.3. Recognition of phospho-serine/threonine modifications .....	74

1.9.4. Recognition of unmodified amino acid residues in histones.....	75
1.10. The PHD Finger: A Versatile Protein-Interaction Domain.....	76
1.10.1. PHD fingers as chromatin binding modules .....	77
1.10.2. Tandem PHD fingers as histone acetyl-lysine sensors.....	83
1.10.3. PHD fingers as protein-protein interaction domains.....	85
1.10.4. PHD E3 ligase activity .....	86
1.10.5. PHD fingers as nuclear phosphoinositide receptors.....	87
1.11. PHD Fingers as Epigenetic Effectors and Potential Drug Targets.....	88
1.12. Lysine Methyltransferases: Structure and Function of SUV39H1 .....	92
1.13. Structure and Function of HP1 .....	98
1.14. Project Aims and Objectives .....	104
<b>Chapter Two: Materials and Methods .....</b>	<b>106</b>
2.1. Sources of Materials.....	107
2.1.1. General suppliers.....	107
2.1.2. Bacterial reagents .....	107
2.1.3. Molecular biology reagents.....	107
2.1.4. Tissue culture reagents .....	108
2.1.5. Biochemical reagents .....	108
2.2. Bacterial preparation and culture .....	109
2.2.1. Composition of solutions and media used in bacterial procedures .....	109
2.2.2. Culture of <i>Escherichia coli</i> strains .....	109
2.2.3. Preparation of <i>Escherichia coli</i> competent cells .....	110
2.2.4. Transformation into <i>Escherichia coli</i> cells .....	110
2.2.5. Long-term storage of bacterial cultures .....	111
2.3. Molecular biology techniques .....	111
2.3.1. Composition of solutions used in molecular biology methods .....	111
2.3.2. Small-scale plasmid purification from bacterial cells .....	112
2.3.3. Large-scale plasmid purification from bacterial cells .....	112
2.3.4. Phenol/chloroform extraction and ethanol precipitation of DNA .....	112
2.3.5. Ethanol precipitation .....	113
2.3.6. Spectrophotometric quantification .....	113

2.3.7. Sequencing of plasmid DNA .....	114
2.3.8. Agarose gel electrophoresis .....	114
2.3.9. Isolation and purification of DNA from agarose gels .....	115
2.3.10. Oligonucleotides .....	115
2.2.11. Polymerase chain reaction (PCR) .....	115
2.3.12. Generation of mutations/deletions by recombinant PCR .....	116
2.3.13. Restriction digests .....	118
2.3.14. Phosphatase treatment .....	118
2.3.15. Ligation reactions .....	119
2.3.16. Colony PCR screening .....	119
2.4 Cell culture .....	120
2.4.1. Maintenance of adherent cell lines .....	120
2.4.2. Maintenance of suspension cell lines .....	120
2.4.3. Cryopreservation of cell lines .....	121
2.4.4. Resuscitation of cell lines .....	122
2.4.5. Calcium phosphate-mediated transfection of adherent cells .....	122
2.5. Biochemical techniques .....	123
2.5.1. Composition of solutions used in biochemical methods .....	123
2.5.2. Preparation of whole-cell extracts .....	125
2.5.3. Preparation of cytoplasmic and nuclear extracts .....	126
2.5.4. Protein concentration determination .....	126
2.5.5. Sodium-dodecyl-sulphate polyacrylamide gel electrophoresis (SDS-PAGE) ...	127
2.5.6. Staining of SDS-polyacrylamide gels .....	127
2.5.7. Western blotting and immunodetection .....	128
2.5.8. <i>In vitro</i> transcription and translation (IVT) of expression vector-encoded cDNA .....	129
2.5.9. Fixing and amplifying radioactive gels .....	130
2.5.10. Exposure of radioactive gels to X-ray film .....	130
2.5.11. Co-immunoprecipitation .....	130
2.5.12. Chromatin immunoprecipitation (ChIP) assays .....	131
2.5.13. Small-scale bacterial expression of GST-tagged proteins .....	133
2.5.14. Large-scale bacterial expression of GST-tagged proteins .....	133
2.5.15. Purification of GST-tagged proteins .....	134
2.5.16. Elution and preparation of GST-tagged protein .....	134

2.5.17. GST-pulldown of histone extracts .....	135
2.5.18. GST-pulldown of mammalian expressed protein.....	136
2.5.19. Peptide binding assay .....	136
2.5.20. Luciferase reporter assays .....	137
2.5.21. Indirect immunofluorescence .....	139
2.2.22. Fluorescence polarisation.....	140
2.5.23. <i>In vitro</i> histone acetyltransferase (HAT) assay .....	140

## **Chapter Three: Results ..... 142**

### **Recognition of histone PTMs by the double PHD finger domain of MOZ**

3.1. Introduction .....	143
3.2. Sequence alignments of MOZ and other PHD fingers.....	145
3.3. The N-terminal domains of MOZ bind to histone H3.....	151
3.4. Binding of MOZ to histone H3 requires the integrity of the DPF .....	159
3.5. Histone H3 and H4 PTMs alter the interaction with MOZ DPF.....	162
3.6. Both PHD1 and PHD2 are required to establish H3 PTM binding preference .....	172
3.7. MOZ is excluded from chromatin containing H3K4me3 .....	177
3.8. MOZ colocalises with chromatin enriched in H3K9me3 and acetylated histone H3 PTMs .....	179
3.9. The MOZ DPF facilitates simultaneous reading of H3 and H4 PTMs .....	184
3.10. The MORF DPF exhibits a similar binding pattern to MOZ DPF.....	189
3.11. Quantitative analysis of MOZ DPF binding to unmodified or acetylated H3/H4 peptides.....	192
3.12. Comparison of MOZ DPF to DPF3 .....	196
3.13. Structural characterisation of the MOZ DPF .....	201
3.14. Summary .....	208



<b>Chapter Four: Results .....</b>	<b>218</b>
<b>The MOZ DPF domain exhibits a H3-specific acetyltransferase activity</b>	
4.1. Introduction .....	219
4.2. The MOZ DPF contains an intrinsic HAT activity .....	220
4.3. Trimethylation of histone H3K4 blocks acetylation of histone H3.....	229
4.4. Cooperative histone acetylation by the DPF+MYST module.....	236
4.6. Acetylation of histone H3 by MOZ DPF promotes its interaction with H3..	241
4.7. The MOZ DPF domain is required for AML-1 mediated transcription.....	246
4.8. Mutation of MOZ DPF does not affect its subcellular localisation .....	250
4.9. Summary .....	252
 <b>Chapter Five: Results.....</b>	 <b>260</b>
<b>MOZ interacts with other chromatin modulators</b>	
5.1. Introduction .....	261
5.2. MOZ interacts with HP1 .....	262
5.3. MOZ colocalises with HP1 $\gamma$ .....	268
5.4. Differential effects of MOZ and HP1 $\gamma$ on AML1-mediated transcription....	271
5.5. MOZ and SUV39H1 interact <i>in vitro</i> .....	275
5.6. MOZ interacts with SUV39H1 <i>in vivo</i> .....	277
5.7. SUV39H1 interacts with two regions of MOZ .....	280
5.8. MOZ does not rescue SUV39H1-mediated repression of AML1 transactivation .....	283
5.9. Summary .....	288
 <b>Chapter Six: Concluding Remarks.....</b>	 <b>296</b>
 <b>Appendix .....</b>	 <b>303</b>
 <b>References .....</b>	 <b>312</b>

## List of Figures

### Chapter One: Introduction

Figure 1.1	Model of PIC formation during transcriptional initiation	4
Figure 1.2	X-Ray structure of the nucleosome core partical at 1.9Å resolution	12
Figure 1.3	Overview of the different classes of modification identified on histones	18
Figure 1.4	Modular structure of lysine acetyltransferases	38
Figure 1.5	Modular structure of MOZ and its interacting partners	45
Figure 1.6	Schematics of leukaemia-associated MOZ fusion proteins	56
Figure 1.7	Chemical structures of known KAT inhibitors	63
Figure 1.8	Chromatin-binding domains: readers and interpreters of the histone code	67
Figure 1.9	Chromatin-binding modules: structural interpretation of the histone code	72
Figure 1.10	Recognition of the H3K4me3 modification by the PlantHomeo domain	79
Figure 1.11	Recognition of the H3K4me0 modification by the BHC80 PHD finger	82
Figure 1.12	Schematic representation of the positive and negative regulation of SUV39H1 activity	97
Figure 1.13	Schematic representation of HP1 paralogs in <i>Drosophila</i>	100

### Chapter Two: Materials and Methods

Figure 2.1	Generation of mutations/deletions by recombinant PCR	117
------------	--	-----

### Chapter Three: Results

Figure 3.1	Sequence alignment of PHD fingers that recognise histone H3 in a H3K4 methylation dependent manner	148
------------	--	-----

Figure 3.2	Sequence alignment of PHD fingers that bind unmodified histone H3	<b>149</b>
Figure 3.3	Sequence alignment with the DPF family	<b>150</b>
Figure 3.4	Creation of GST-fusion and FLAG-constructs used in this study	<b>152</b>
Figure 3.5	Establishing optimal conditions for the GST pulldown assay	<b>154</b>
Figure 3.6	The N-terminal domains of MOZ bind to Histone H3	<b>155</b>
Figure 3.7	Dose-dependent binding of MOZ N-terminal domains to Histone H3	<b>158</b>
Figure 3.8	Binding of MOZ to Histone H3 requires the integrity of the DPF	<b>161</b>
Figure 3.9	Binding of MOZ to Histone H3 PTMs requires the integrity of the DPF	<b>163</b>
Figure 3.10	Validation of the modification status of biotinylated histone peptides used in binding studies	<b>165</b>
Figure 3.11	Binding preference of the MOZ DPF: An intolerance of H3K4me3 and H4Ac	<b>168</b>
Figure 3.12	Core histone extracts are hyperacetylated on histone H4	<b>169</b>
Figure 3.13	Binding specificity of the full-length MOZ protein: An intolerance of H3K4me3 and H4Ac	<b>171</b>
Figure 3.14	Binding specificity of the individual PHD motifs of the MOZ DPF	<b>175</b>
Figure 3.15	Both PHD1 and PHD2 are required to establish H3 PTM binding preference	<b>176</b>
Figure 3.16	MOZ is excluded from chromatin containing H3K4me3	<b>178</b>
Figure 3.17	Exclusion of MOZ from H3K4me3 enriched chromatin	<b>182</b>
Figure 3.18	Colocalisation of MOZ with acetylated H3 enriched chromatin	<b>183</b>
Figure 3.19	The MOZ DPF facilitates simultaneous reading of H3 and H4 PTMs	<b>186</b>
Figure 3.20	Dose-dependent disruption of MOZ-DPF binding to H3Kac by H4Kac	<b>187</b>
Figure 3.21	MORF DPF exhibits a similar histone binding pattern as MOZ DPF	<b>191</b>

Figure 3.22	Quantitative analysis of MOZ DPF binding to site-specific histone H3 and H4 PTMs	195
Figure 3.23	Conservation of surface exposed residues based on the hDPF3b NMR structure	199
Figure 3.24	Alignment of DPF3b and the structural homology model of MOZ DPF	200
Figure 3.25	Crystallisation of the MOZ DPF domain	202
Figure 3.26	Two-dimensional structure of the tandem PHD finger of human MOZ	205
Figure 3.27	Structural mechanism of H3K4 engagement by human MOZ DPF	206
Figure 3.28	Comparison of MOZ and DPF3b tandem PHD fingers	207
Figure 3.29	Schematic model depicting how combinations of H3 and H4 PTMs influence the engagement of MOZ DPF to chromatin	213
Figure 3.30	Schematic model depicting the possible sequence of events involved in MOZ/MLL/WDR5 complex histone acetylation and methylation	217

## Chapter Four: Results

Figure 4.1	Purification of the MOZ DPF and MYST domains	222
Figure 4.2	The MOZ DPF contains an intrinsic HAT activity	224
Figure 4.3	HAT activity is a property of the intact DPF module	227
Figure 4.4	Acetylation activity of the full-length MOZ protein	228
Figure 4.5	Trimethylation of histone H3K4 blocks the acetylation of histone H3	231
Figure 4.6	MOZ DPF primarily acetylates lysine 9 of histone H3	232
Figure 4.7	Acetylation of histone H4 blocks further acetylation by MOZ	235
Figure 4.8	The DPF domain enhances the HAT activity of MOZ	238
Figure 4.9	The DPF domain enhances the rate of histone acetylation by MOZ	240

Figure 4.10	Acetylation of histone H3 by MOZ DPF promotes its interaction with H3	243
Figure 4.11	H4 acetylation does not affect the acetylation of histone H3	245
Figure 4.12	MOZ DPF mutants inhibit transcription activation of the AML1 reporter	249
Figure 4.13	Subcellular localisation of exogenously expressed MOZ proteins	251

## Chapter Five: Results

Figure 5.1	HP1 $\alpha/\beta$ interaction with the N-terminus of MOZ <i>in vitro</i>	264
Figure 5.2	MOZ interacts with HP1 $\gamma$ <i>in vivo</i>	266
Figure 5.3	Colocalisation of MOZ with H3K9me3 and HP1 $\gamma$	270
Figure 5.4	Over-expression of MOZ does not rescue HP1 $\gamma$ -mediated repression at the AML1 promoter	274
Figure 5.5	MOZ interacts with the lysine methyltransferase SUV39H1 <i>in vitro</i>	276
Figure 5.6	MOZ interacts with the lysine methyltransferase SUV39H1 <i>in vivo</i>	279
Figure 5.7	SUV39H1 interacts with two regions of MOZ	282
Figure 5.8	SUV39H1 represses the transcriptional activity of AML1	286
Figure 5.9	Over-expression of MOZ does not rescue SUV39H1-mediated repression at the AML1 promoter	287
Figure 5.10	Schematic model illustrating possible events at the AML1 reporter	291
Figure 5.11	Model showing the hypothesised recruitment of transcriptional modulators by MOZ to chromatin	294

## Chapter Six: Concluding Remarks

Figure 6.1	Schematic model depicting the recruitment and release signals for MOZ to and from chromatin	298
Figure 6.2	Schematic model depicting the possible sequence of events involved in MOZ function	302

## List of Tables

### Chapter Two: Materials and Methods

Table 2.1	Solution composition for resolving and stacking gels for SDS-PAGE	128
-----------	---	-----

### Appendix

Table A.1	Oligonucleotide Sequences Used for Cloning Purposes	304
Table A.2	Oligonucleotide Sequences Used for Site-Directed Mutagenesis	307
Table A.3	Origin of Plasmids	308
Table A.4	Plasmids Created by Cloning or Site-directed Mutagenesis	309
Table A.5	Origin and Dilutions of Antibodies	311

## Abbreviations

<b>Acetyl-CoA</b>	Acetyl-CoenzymeA
<b>ACF</b>	ATP-dependent Chromatin assembly and remodelling Factor
<b>AdoMet</b>	S-adenosylmethioine
<b>AD</b>	Activation domain
<b>ADP</b>	Adenosine Di-Phosphate
<b>ADR</b>	Adriamycin
<b>AIRE</b>	Autoimmune Regulator
<b>ALL</b>	Acute Lymphoblastic Leukaemia
<b>AML</b>	Acute Myeloid Leukaemia
<b>AML1</b>	Acute Myeloid Leukaemia Protein 1
<b>ATP</b>	Adenosine Tri-Phosphate
<b>ATR-X</b>	Alpha-Thalassaemia Mental Retardation, X-linked
<b>BD</b>	Bromodomain
<b>β-GAL</b>	β-Galactosidase
<b>BHC80</b>	BRAF35-HDAC 80
<b>bp</b>	Base pairs
<b>BPTF</b>	Bromodomain and PHD domain Transcription Factor
<b>BRPF1</b>	Bromodomain- and PHD Finger-containing Protein 1
<b>BSA</b>	Bovine Serum Albumin
<b>CARM1</b>	Coactivator-associated arginine methyltransferase 1
<b>CBP</b>	CREB Binding Protein
<b>CD</b>	Chromodomain
<b>CGI</b>	CpG Islands
<b>CHD</b>	Chromodomain, Helicase, DNA binding
<b>ChIP</b>	Chromatin Immunoprecipitation
<b>CID</b>	CBP Interacting Domain
<b>C/N</b>	Cytosol/Nuclear
<b>CNC</b>	Cranial Neural Crest Cells
<b>CPI-2</b>	Calpain/cathepsin Inhibitor 2
<b>CREB</b>	cAMP Response Element Binding Protein
<b>CSD</b>	ChromoShadow Domain
<b>CSF1R</b>	Macrophage Colony-Stimulating Factor Receptor
<b>CTIP2</b>	COUP-TF-Interacting Protein 2
<b>DBC1</b>	Deleted in Breast Cancer 1
<b>DDR</b>	DNA Damage Response
<b>DMEM</b>	Dulbecco's Modified Eagle Medium
<b>DMSO</b>	Dimethyl Sulfoxide
<b>DNA</b>	Deoxyribo Nucleic Acid
<b>DPF</b>	Double PHD Finger
<b>DSIF</b>	DRB- Sensitivity Inducing Factor
<b>DTT</b>	Dithiothreitol
<b>EB</b>	Elution Buffer
<b>ECL</b>	Enhanced Chemiluminescence

<b>EDTA</b>	Ethylenediaminetetraacetic acid
<b>EGTA</b>	Ethylene Glycol Tetraacetic Acid
<b>EMBL</b>	European Molecular Biology Laboratory
<b>ENL</b>	Eleven Nineteen Leukaemia
<b>E(Z)</b>	Enhancer of Zeste
<b>FACT</b>	Facilitates Chromatin Transcription
<b>FCS</b>	Fetal Calf Serum
<b>GAG</b>	Glycosaminoglycan
<b>GFP</b>	Green Fluorescent Protein
<b>GNAT</b>	Gcn5-related N-acetyltransferase
<b>GST</b>	Glutathione-S-Transferase
<b>GTF</b>	General Transcription Factor
<b>H1,2A,2B,3,4</b>	Histone 1,2A,2B,3,4
<b>HAT</b>	Histone Acetyltransferase
<b>HBO1</b>	Histone Acetyltransferase Binding to ORC-1
<b>HBS</b>	HEPES Buffered Saline
<b>HCS</b>	HoloCarboxylase Synthetase
<b>HDAC</b>	Histone Deacetyltransferase
<b>HEPES</b>	4-(2-HydroxyEthyl)-1-PiperazineEthaneSulfonic acid
<b>HF</b>	High-Fidelity
<b>HMT</b>	Histone MethylTransferase
<b>Hox</b>	Homeobox
<b>HP1</b>	Heterochromatin Protein 1
<b>HRP</b>	Horseradish Peroxidase
<b>HSC</b>	Hematopoietic Stem Cells
<b>HSPG</b>	Heparan Sulphate Proteoglycans
<b>IPTG</b>	Isopropyl $\beta$ -D-1-thiogalactopyranoside
<b>ING</b>	Inhibitor of Growth
<b>INO80</b>	Inositol requiring 80
<b>IP</b>	Immunoprecipitation
<b>ISWI</b>	Imitation SWI
<b>IVT</b>	<i>In Vitro</i> Translation
<b>KAP1</b>	KRAB-ZPF Associated Protein 1
<b>KAT</b>	Lysine Acetyltransferase
<b>KDAC</b>	Lysine
<b>KIX</b>	Kinase-Inducible Domain
<b>KMT</b>	Lysine Methyltransferase
<b>LB</b>	Luria-Bertani
<b>LSD1</b>	Lysine-Specific Demethylase 1
<b>mRNA</b>	Messenger RNA
<b>MART</b>	Mono-ADP-Ribosyltransferase
<b>MBT</b>	Malignant Brain Tumour
<b>MLL</b>	Mixed Lineage Leukaemia
<b>MOF</b>	Males Absent On the First
<b>MORF</b>	Monocytic Leukaemia Zinc-finger Related Factor
<b>MOZ</b>	Monocytic Leukaemia Zinc-finger protein



<b>MPO</b>	Myeloperoxidase
<b>MYST</b>	MOZ, Ybf2/Sas3, Sas2, TIP60
<b>NAD</b>	ADP-Ribosyl Nicotinamide
<b>NAP1</b>	Nucleosome Assembly Protein 1
<b>NCP</b>	Nucleosome Core Particle
<b>NEB</b>	New England Biolabs
<b>NELF</b>	Negative Elongation Factor
<b>NLS</b>	Nuclear Localisation Signal
<b>NMR</b>	Nuclear Magnetic Resonance
<b>NUP98</b>	Nucleoporin 98
<b>NURF</b>	Nucleosome Remodelling Factor
<b>OD</b>	Optical Density
<b>PADI4</b>	Peptidyl Arginine Deiminase 4
<b>PARP</b>	Poly-ADP-Ribose Polymerases
<b>PBS</b>	Phosphate Buffer Saline
<b>PCAF</b>	p300/CBP Associated Factor
<b>PcG</b>	Polycomb Group
<b>PCL</b>	Polycomblike
<b>PCR</b>	Polymerase Chain Reaction
<b>PDB</b>	Protein Data Bank
<b>PEV</b>	Position Effect Variegation
<b>PHD</b>	Plant HomeoDomain
<b>PIC</b>	Pre-Initiation Complex
<b>Pin</b>	Parvulin
<b>PML</b>	Promyelocytic leukaemia nuclear bodies
<b>PMSF</b>	PhenylMethaneSulfonylFluoride
<b>PRMT</b>	Protein arginine <i>N</i> -methyltransferases
<b>PtdInsPs</b>	Phosphoinositides
<b>PTM</b>	Post-Translational Modification
<b>RAG2</b>	Recombination Activating Gene 2
<b>RAR<math>\beta</math>2</b>	Retinoic Acid Receptor $\beta$ 2
<b>RING</b>	Really Interesting New Gene
<b>RIPA</b>	RadioImmunoprecipitation Assay buffer
<b>RLU</b>	Relative Luciferase Units
<b>RNA</b>	Ribo Nucleic Acid
<b>rRNA</b>	Ribosomal RNA
<b>RNAP</b>	RNA Polymerase
<b>RPMI</b>	Roswell Park Memorial Institute medium
<b>RT</b>	Room Temperature
<b>RTS</b>	Rubinstein-Taybi Syndrome
<b>S phase</b>	Synthesis phase
<b>SAGA</b>	Spt-Ada-Gcn5 Acetyltransferase
<b>SAH</b>	S-Adenosylhomocystein
<b>SDS</b>	Sodium Dodecyl Sulphate
<b>SDS-PAGE</b>	Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis

<b>SET</b>	Su(var)3-9, Enhancer of Zeste, Trithorax
<b>SID</b>	Steroid Receptor Coactivator 1 Interaction Domain
<b>SIM</b>	SUMO Interacting Motif
<b>siRNA</b>	Short interfering RNA
<b>Sir</b>	Sirtuin
<b>snRNA</b>	Small nuclear RNA
<b>SUMO</b>	Small Ubiquitin-related Modifier
<b>SUV39H1</b>	Suppressor of Variegation 39
<b>SWI/SNF</b>	Switching defective/Sucrose Nonfermenting
<b>tRNA</b>	Transfer RNA
<b>TAF</b>	TBP Associated Factors
<b>TAQ</b>	Thermus Aquaticus
<b>TBE</b>	Tris-borate-EDTA
<b>TBP</b>	TATA-Binding Protein
<b>T-B-SCID</b>	Severe Combined Immunodeficiency
<b>TE</b>	Tris- EDTA
<b>TSAP</b>	Thermosensitive Alkaline Phosphatase
<b>TIF2</b>	Transcriptional Intermediary Factor 2
<b>TIP60</b>	Tat Interacting Protein, 60kDa
<b>TRIM24</b>	Tripartite motif-containing protein 24
<b>TSS</b>	Transcription Start Site
<b>UAS</b>	Upstream Activating Sequence
<b>Ubc9</b>	Ubiquitin-like protein SUMO-1 Conjugating enzyme
<b>UV</b>	Ultraviolet
<b>WDR5</b>	WD-Repeat Domain 5
<b>WT</b>	Wild-type

# **Chapter One: Introduction**

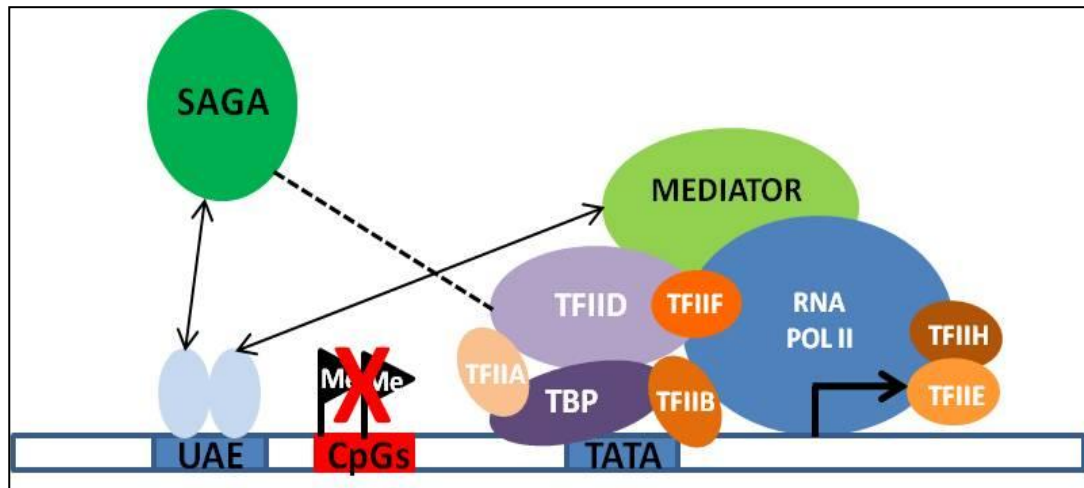
### 1.1. Transcription

The stringent control of gene expression during development, differentiation, and the maintenance of cellular function, is ultimately essential for the survival and evolution of complex multi-cellular organisms. Transcription is one of the first steps in gene expression and this involves molecular events to copy the genes encoded in eukaryotic chromosomes into RNA. Indeed, many biological processes in living cells and organisms are regulated primarily at the level of transcription.

The transcriptional process relies upon a basal machinery of evolutionary conserved transcription factors and polymerases, which collectively reconstitute a transcription initiation complex. Within the eukaryotic cell nucleus there are five functionally distinct RNA polymerases (RNAP), designated RNA polymerase I, II, III, IV and V. RNAP I and its transcription machinery function in the nucleolus to synthesise ribosomal (r)RNA. rRNA biogenesis regulates ribosome production that in succession guides cell growth and proliferation (Drygin et al., 2010; Grummt, 1999). RNAP III is required for the production of short RNA molecules, including transfer (t)RNA and rRNA which are implicated in protein synthesis and RNA maturation (Canella et al., 2010; Willis, 1993). RNAP IV and V synthesise short interfering (si)RNA that are required for heterochromatin formation in plants (Herr et al., 2005; Wierzbicki et al., 2009). Finally, RNAP II is responsible for the synthesis of all eukaryotic messenger (m)RNA and most small nuclear (sn)RNA and microRNA, which are respectively utilised as templates for protein synthesis, RNA processing and the regulation of gene activity to name but a few. The functional importance of

RNAP II is further reflected by the fact that it is the most extensively studied of this class of enzymes.

As discussed, the initiation of RNA synthesis has emerged as a primary event in determining cellular gene expression levels (Tjian, 1996), and therefore it is not surprising that RNAP II is the end target of a vast array of signal transduction pathways. It is this elaborate regulation of RNAP II transcription that underlies the control of cellular development and differentiation (Hahn, 2004). RNAPII is unable to activate transcription on its own: the enzyme is capable of uncoiling and recoiling DNA, and the synthesis of RNA, but is incapable of recognising and binding a promoter sequence (Kornberg, 2007). This role requires the participation of the general transcription factors (GTFs). Many groups contributed towards the isolation of the RNAP II transcription proteins and these findings were critical for the solution of the structure and regulation of the RNAP II machinery (Thomas and Chiang, 2006). Further, recent advances in structure determination have led to an increased understanding of the precise assembly of transcription proteins at the promoter and the molecular events that occur during the initiation of transcription (Hahn, 2004).



**Figure 1.1. Model of PIC formation during transcriptional initiation.** Upon targeting to the Upstream Activation Sequence (UAS), activators recruit coactivators (including SAGA, Mediator, Swi/Snf). This recruitment further augments the binding of activators, triggers hyperacetylation of histones and the subsequent remodelling of nucleosomes. Unmethylated CpG Islands (CGIs) also promote a permissive chromatin state. The resulting exposure of the entire core promoter and the SAGA/Mediator complexes facilitate the assembly of GTFs and RNAP II to form the PIC and initiate transcription. (Figure adapted and updated from Hahn, 2004).

The high level of control required over transcription is dictated by regions of eukaryotic protein coding genes. Each gene carries a core promoter region and a unique array of proximal and distal regulatory elements, which are recognised by sequence specific DNA binding factors critical for the activation or repression of transcription initiation. For illustration, activators bind to enhancer sequences (e.g. Upstream Activation Sequences - UAS), located near to or far removed from gene promoters, triggering a cascade of recruitment of coactivator complexes; including histone modification enzymes, chromatin remodelers and Mediator (Figure 1.1).

Many genes transcribed by RNAP II have well defined nucleosomal positioning at promoters, thus SAGA histone acetyltransferase activity, histone acetylation and the Swi/Snf remodeler act prior to Pre-Initiation Complex (PIC) formation, in order to overcome this nucleosomal barrier and make DNA more accessible (Reinke et al., 2001). Mediator, a large multi-subunit complex conserved in all eukaryotes, interacts directly with enhancer bound activator proteins and, consequently, contacts RNAP II and GTFs at the promoter to stimulate the initiation of transcription (Kornberg, 2005). Mediator is not only the basis for regulated transcription at almost all RNAP II promoters; but is comparable with RNAP II in its importance for transcription. Previously, Mediator has been viewed solely as a coactivator in function, however it is now evident that it can also function as a corepressor and GTF. Mediator can be viewed as a signal processor, relaying both positive and negative regulatory information from enhancers to promoters (Takagi and Kornberg, 2006).

The discovery of CpG islands (CGIs) added greater complexity to our understanding of transcriptional control. The mammalian genome is punctuated with a high frequency of DNA sequences rich in G/C base composition and a high density of CpG dinucleotides (Figure 1.1). CGIs encompass the transcription start site (TSS) and promoter regions of 60-70% of human genes (Illingworth and Bird, 2009) and often lack DNA methylation. They also coincide with sites of histone H3 Lysine 4 trimethylation and colocalise with RNAP II (Thomson et al., 2010). These elements annotate active gene promoters but are refractory to epigenetic silencing by DNA methylation. The link between unmethylated CGIs and transcriptional activation has long been observed, however it has only recently been shown that the underlying sequence of CGIs genetically influences the chromatin modification status through

the recruitment of protein mediators (CpG binding proteins). It has thereby been proposed that the DNA sequence at CGIs simplifies genome function by manipulating chromatin structure directly (Thomson et al., 2010).

After specific positive regulatory events of activator recruitment and nucleosomal rearrangement, promoter DNA is made transiently available for the interaction with the transcription machinery. Typically, the RNAP II transcription cycle begins with the binding of gene-specific regulatory factors at sequence elements within the core promoter; these elements include the TATA box. Core promoter recognition is required for the correct positioning and assemblage of RNAP II and GTFs to form the PIC (Figure 1.1).

TFIID consists of the TATA-box binding protein (TBP), which allows attachment at the promoter sequence and approximately 14 TBP associated factors (TAFs) that modulate promoter recognition. The sequential binding of the TFIIA and TFIIB transcription factors further stabilise the TFIID-promoter attachment. In a stepwise manner, the pre-formed RNAPII-TFIIF complex, TFIIE and TFIIH bind. TFIIH possesses DNA helicase activity, which exposes the template strand at the transcription start point. The subunits of TFIIH are also important in nucleotide base excision repair, and ultimately promoter clearance that results in the start of the elongation phase (Zhovmer et al., 2010). The assembly of the complete closed PIC is not in an active conformation to begin transcription, it is not until a clear conformational change occurs, in which 11-15 base pairs (bp) of DNA surrounding the TSS is melted and the template strand is exposed to the active site of RNAP II to form the open complex (Hahn, 2004). Transcriptional initiation is set in motion with the formation of the first phosphodiester bond of RNA.



The regulation of transcription is a multi-step process, with stringent control being exerted at the elongation and termination phases as well as the initiation stage. RNAP II clears the promoter and enters early elongation phase, after releasing contacts with the core promoter and the rest of the transcription machinery. Studies including ChIP-chip and permanganate footprinting, were used to analyse the distribution of transcriptionally engaged RNAP II across the genome (Core and Lis, 2008). Results revealed an enriched RNAP II density positioned within the 5' body of the gene. This concept, named promoter-proximal pausing describes how transcriptionally engaged RNAP II transcribes 20-50bp downstream of the TSS, producing an RNA and pauses (Core and Lis, 2008). This pause correlates with the action of the pausing factors DSIF (DRB- sensitivity inducing factor) and NELF (negative elongation factor), which aid in stabilising RNAP II in the paused form (Yamaguchi et al., 1999). Activator proteins lead to the attraction of factors that promote RNA chain synthesis, processing, export and chromatin modification to release RNAP II from its paused state. RNAP II and its associated pausing factors are hyperphosphorylated, allowing the enzyme to escape the pause and either enters productive elongation or RNA synthesis terminates. Promoter-proximal pausing is a post-recruitment regulation stage that allows promoters to be held in an open and accessible configuration, ensuring the rapid and synchronous transition into productive elongation and activation of gene expression upon stimuli from additional factors (Fuda et al., 2009). More recently, the deep sequencing of the 3' ends of nascent RNA transcripts, associated with RNA polymerase, has allowed the visualisation of transcription at a nucleotide resolution. Nascent RNA sequencing has revealed that pervasive polymerase pausing and backtracking is widespread, both at

human promoters and throughout the entire body of the transcript (Core et al., 2008; Churchman and Weissman, 2011).

During elongation, nucleosomes are displaced and redeposited to allow the passage for RNAP II to transcribe throughout the entire gene body. After the RNAP II complex has transcribed the gene, transcription is terminated, RNAP II removed from the DNA and the RNA released. Freed RNAP II can then be recycled and reused in subsequent rounds of transcriptional initiation (Fuda et al., 2009).

Current models establish that polymerases are recruited to their target genes wherever their location may be. However, novel revelations in the field of gene regulation are emerging, providing evidence for alternative mechanisms for polymerase action. Cook and colleagues have provided evidence, through mapping with nucleases and chromosome conformation capture, that active RNA polymerases and associated factors cluster into ‘factories’. Within these factories many of the functions needed to produce mature RNA transcripts are carried out (Faro-Trindade and Cook, 2006). This clustering into distinct cellular compartments ensures efficient interaction between molecules, leading consequently to increased reaction rates and enhanced regulation. Cook’s current model states that genes, as part of chromatin loops, must diffuse to the appropriate transcription factory before expression can take place (Faro-Trindade and Cook, 2006). Active, hyperphosphorylated polymerases are tethered to these factories and act both to maintain the structural integrity of the chromatin loops and as motors, reeling in their template DNA (Cook, 1999). Recently, it has been shown that factories specialise in the production of a specific type of transcript depending on promoter type, whether or not the gene contains an intron (Xu and Cook, 2008) and whether the gene products are expressed co-

ordinately (Carter et al., 2008): however, how many different types of factory there are remains to be established. In the absence of transcription these RNAPII foci remain, indicating they are not simple accumulations of RNAPII on transcribed genes but exist as independent nuclear subcompartments (Mitchell and Fraser, 2008).

Our knowledge of nuclear organisation and transcriptional regulation is ever expanding; however it is important to highlight work carried out over the past few years outlining the biological importance of the main interfaces between transcription and mechanisms maintaining genome integrity. High levels of transcription correlate with common fragile site breakage within chromosomal regions (Yu et al., 2000), mutagenesis within a regulatory element or coding region of an individual or group of genes (Epstein, 2009; Saxowsky and Doetsch, 2006), and elevated levels of DNA recombination (Aguilera, 2002); thus fundamentally impinging on genome stability. In order to maintain genome integrity, an invariable cross-talk is required between transcription and other basic cellular processes, including DNA damage and repair pathways and chromatin remodelling (Svejstrup, 2010). Certification for the biological significance of this cross-talk comes from the finding that mutations in the genes encoding those proteins that mediate cross-talk can cause severe human diseases; such as the skin malignancy Xeroderma pigmentosum or Cockayne syndrome which is characterised by growth retardation, premature aging and an impairment to nervous system development (Nospikel, 2008). This relatively poorly understood area of DNA research is still in its infancy, however researchers are beginning to consider and understand how processes have evolved to guarantee that the genome destabilising effect of transcription is minimised.

### 1.2. Chromatin Structure and Dynamics Govern Cellular Processes

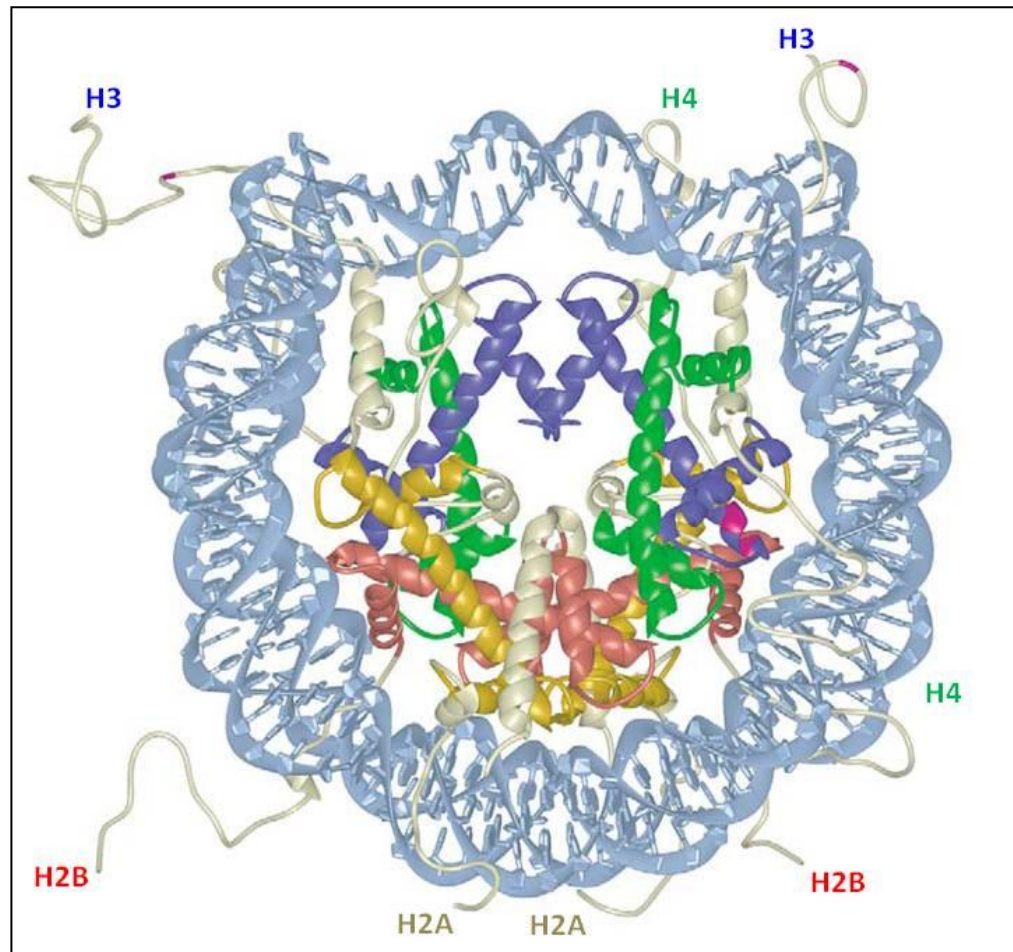
All eukaryotic organisms have elaborate ways of organising and packaging approximately two metres of DNA within a volume of several microns (Peterson and Laniel, 2004). This hierarchical scheme of compression is performed by nuclear proteins that successively coil and fold the DNA into higher order structures. This complex consisting of DNA, nuclear histones and an array of different chromosomal proteins is known as chromatin.

Chromatin is a structural polymer that not only solves the basic packaging problem, but also provides a dynamic platform that assists and dictates the progression of nuclear processes; such as transcription, replication, cell-cycle progression, DNA repair and recombination. The basic repeating structure of this nucleoprotein complex is the evolutionary conserved nucleosome. The nucleosome core particle (NCP) consists of 147bp of DNA wrapped nearly twice around a histone octamer core, comprising two copies each of the histones H2A, H2B, H3 and H4 (Luger, 2003) (Figure 1.2).

The S-phase synthesised core histones are predominantly globular in structure, except for an amino-terminal 20-35 residue segment that protrudes from the surface of the nucleosome (Peterson and Laniel, 2004). The central globular domain mediates both inter-nucleosomal, histone-histone interactions and the direct contact with the DNA backbone, allowing consequent organisation of the two wraps of nucleosomal DNA. These multiple interactions make the nucleosome one of the most stable protein-DNA complexes under physiological conditions (Li et al., 2007a). The

unstructured N-terminal tails are rich in basic amino acid residues and are required for the ordered condensation of nucleosomal arrays into higher-compacted structures. Histone proteins play structural and functional roles in the regulation of many nuclear processes including transcriptional control. Tight regulation prevails in the form of small molecule, reversible modifications of histones, especially on their N-terminal tails. These modifications serve as an epigenetic code, signalling for changes in chromatin architecture. This topic will be discussed in greater detail in subsequent sections.

In addition to the four core histones constituting the NCP, a fifth histone H1 (linker histone) is present in nuclei. Histone H1 associates with the non-constrained linker DNA within the nucleosome array, offering partial nuclease protection (Woodcock and Ghosh, 2010). However, linker histones are best classified as chromatin architectural proteins; proteins which stabilise and promote higher order structures by converting nucleosomal arrays into stable 30nm structures with enhanced compaction and regularity (Routh et al., 2008). Recently, accumulating evidence portrays that aside from its well-established role in chromatin fibre stabilisation, linker histones can act as recruitment hubs for many non-histone nuclear and cytosolic proteins involved in overlapping processes centred on genomic DNA. Many cases have been documented showing linker histone potential to mediate multiple simultaneous interactions when bound to the nucleosome (McBryant et al., 2010).



**Figure 1.2. X-Ray structure of the nucleosome core particle at 1.9Å resolution.** Figure taken from (Luger, 2003). Copyright clearance obtained from Elsevier publishing (licence number: 2798250403120). The nucleosome core particle as viewed down the superhelical axis. The core histone octamer is composed of a central heterotetramer of histones H3 (blue) and H4 (green), flanked by two heterodimers of histones H2A (yellow) and H2B (red). The DNA helix is shown in light blue and protruding histone N-terminal tails in cream.

The nucleosomal array is further condensed by forming a helical coil (30nm fibre) stabilised, in part, by the externally core associated linker histone H1 (Routh et al., 2008) and the histone octamer surface of the nucleosome. Nucleosome arrays are additionally stabilised by the introduction of disulphide crosslinks and show that the chromatin fiber comprises two stacks of nucleosomes, which zig zag back and forth in a two-start helix organisation (Dorigo et al., 2004). *In vitro* experimentation showing the formation of polynucleosomes into a *bona fide* 30nm chromatin secondary structure is generally accepted. However, this level of compaction is the subject of intense debate and remains controversial due to the difficulty in obtaining *in vivo* evidence. Careful examination of isolated 30nm fibres have failed to reveal a consistent, precise arrangement of nucleosomes (Woodcock and Ghosh, 2010) and data from new experimental techniques including chromatin conformation capture and cryo-electron microscopy are beginning to question past evidence of higher order chromatin assemblies based on the 30nm fiber (Fussner et al., 2010).

Due to the lack of clarity and definitive information on chromatin secondary structure, models of higher hierarchical structures become complex. However, current evidence concludes that the final condensation of chromatin involves the *in vivo* formation of 100-400nm interphase fibers (Peterson and Laniel, 2004) or the more highly compacted (10,000-20,000X), consistent manifestation of chromatin; the metaphase chromosome structures (Woodcock and Ghosh, 2010).

In addition to nucleosomes, an array of histone variants and accessory proteins constitute the chromatin fiber and aid the modulation of chromatin structure. Histone variants are non-allelic isoforms of the canonical core histones that have the potential to adjust the stability of nucleosomes and nucleosomal arrays (Park and

Luger, 2008). Further, structural studies disclose that these variants may facilitate intranucleosomal interactions, thereby contributing to the alteration of higher-order chromatin structures (Luger, 2003). Histone variants are mainly distinguished from core histones in that they are expressed outside of S-phase and incorporated into chromatin in a DNA-replication-independent manner (Li et al., 2007a). These variants are not randomly distributed in chromatin but are expressed in developmentally constrained or cell type specific patterns (Verdone et al., 2006). H2A.Z, a variant of the core histone H2A, can be deposited into a nucleosome either through ATP-dependent histone exchange reactions (Mizuguchi et al., 2004) or with the help of the histone chaperone NAP1 (Park et al., 2005). Depending on the position of the nucleosome in which it has been placed, H2A.Z has been shown to exert critical functions to facilitate either transcriptional repression at the heterochromatin boundary, or transcriptional activation at gene promoters. Its importance as a modulator of chromatin structure is reinforced by the fact that it cannot be substituted for by the *bona fide* H2A core histone (Clarkson et al., 1999).

As described previously in section 1.1, the nucleosome and higher order chromatin structures impede the access of enzymes and factors that assist DNA-mediated processes such as transcription, DNA damage and repair. The maintenance of chromatin structure is therefore fundamental for eukaryotic gene regulation. Another group of proteins, the histone chaperones, are mainly responsible for retaining a balance between the assembly and the partial or complete disassembly of nucleosomes in order to sustain genome integrity (Park and Luger, 2008). In general, the chaperones bind histones in order to prevent non-specific interactions between the DNA and histones and promote specific interactions that initiate nucleosome



assembly through ATP-independent reactions (Eitoku et al., 2008). To maintain the balance, histone chaperones also cooperate with chromatin remodelling and modifying enzymes to facilitate the eviction of nucleosomes at promoter regions where DNA-mediated processes can then commence (Park and Luger, 2008).

The literature ubiquitously utilises the terms heterochromatin and euchromatin when discussing chromatin structures. However, it should be documented that these terms refer to the state of compaction and transcriptional potential of chromatin rather than higher order structures *per se* (Woodcock and Ghosh, 2010). Following the completion of mitosis, the compact chromosomal structure required for metaphase is no longer necessary. Therefore, as the cell enters interphase, specific areas of the chromatin decondense yielding two distinct types of chromatin; euchromatin and heterochromatin. Euchromatin is characterised by chromatin domains containing transcriptionally active genes. These regions form chromatin loops that extend away from the compact chromosome territories, towards the interior of the interphase nucleus (Bartova et al., 2008b). Active genes from different chromosome territories located on decondensed chromatin loops are transcribed by the same RNAP II in subnuclear compartments called transcription factories (Faro-Trindade and Cook, 2006). Conversely, heterochromatin is a gene-poor, densely packaged domain localised to the nuclear periphery associated with transcriptional silencing (Bartova et al., 2008a).

Higher order chromatin structures are of great importance for packaging all genetic information within cell nuclei. They also provide a dynamic platform that governs many fundamental nuclear processes, allowing stringent control over the

expression of an individual eukaryotic gene and further, more complex gene networks that require coordinate regulation.

### 1.3. Histone Code Hypothesis

As discussed in preceding sections, eukaryotic gene expression is regulated at the level of chromatin structure. We have discussed that input from gene regulatory sequences, the binding of sequence-specific activators and repressors and additional factors that modify, interact with and remodel chromatin are all required for the correct regulation of gene expression. At the turn of the 21<sup>st</sup> century it emerged that histones, once thought to be static structural elements, were indeed vital and dynamic entities of the machinery accountable for regulating gene transcription and other DNA-templated processes such as replication, repair, recombination and chromosome segregation.

Numerous residues within the N-terminal tails and globular domains of histones are subjected to various post-translational modifications (PTMs); including acetylation, methylation and phosphorylation. These will be discussed in greater detail in a subsequent section (1.4). The huge diversity of modifications and the complex nature of their arrangement, both globally and on individual histone tails, led to the proposal of the concept of a ‘histone code’, where distinct histone modifications act sequentially or in combination to generate signals that are read by other proteins to bring about specific downstream events including the transcriptional state of the gene (Strahl and Allis, 2000). Therefore, in addition to inducing extensive changes to nucleosomal structure, modifications can lead to the alteration of distinct

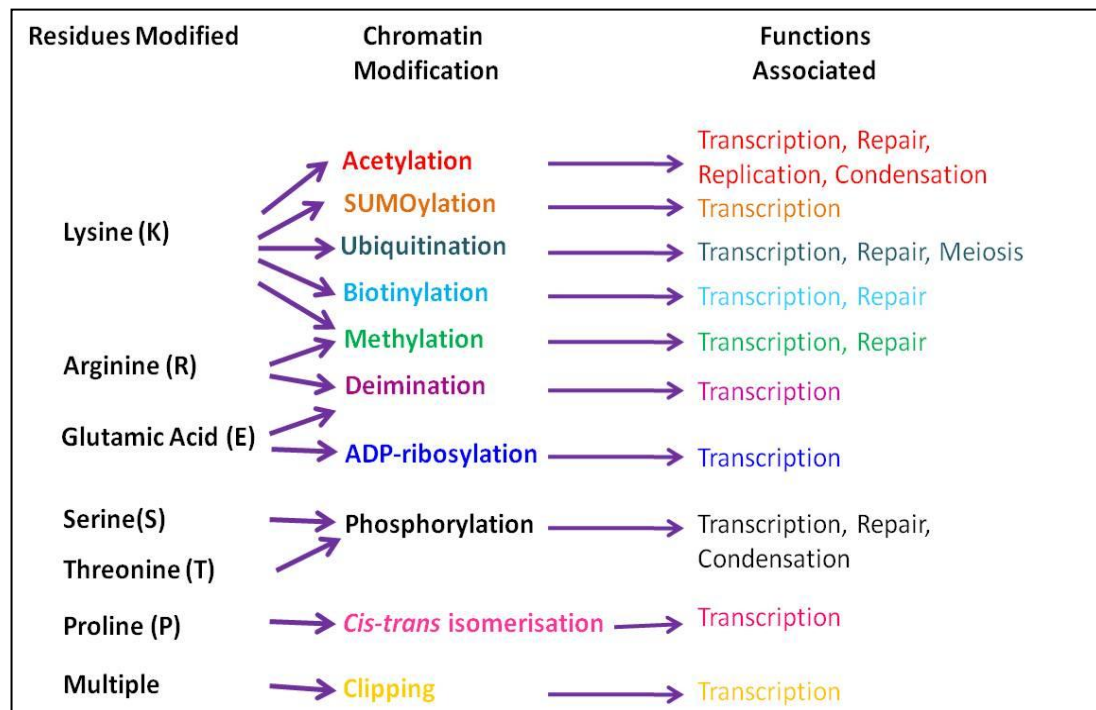
surfaces of the nucleosome. These amendments promote the binding of specific factors to the nucleosome, resulting in the dissociation or recruitment of proteins, dependent on the functional pathway activated.

As research progresses however, it is evident that there are extra levels of intricacy involved in the chromatin signalling pathway, suggesting not a simple code but a complex language determining transcriptional readout. More complex scenarios are arising disclosing crosstalk between multiple modifications on the same or adjacent nucleosomes. In some cases, PTMs acting in combination in both a context and time-dependent manner is crucial for a particular transcriptional readout (Lee et al., 2010). Also, it is becoming clear in recent histone crosstalk research that the order and mechanism of the addition and removal of PTMs are essential for the readout of a gene (Lee et al., 2010). Thus, mapping histone modification trends, without the context of the recruitment, regulation, and interactions of the complexes executing these marks is not sufficient to understand the mechanisms regulating gene expression.

### **1.4. Covalent Modification of Histones and Chromatin Remodeling**

Post-translational modifications are chemical alterations following translation, which ultimately regulate the stability and function of both histone and nonhistone proteins. A prominent attribute of histones is the large number and type of modified residues they possess (Figure 1.3). These modifications are known to function via two characterised mechanisms. The first is associated with the disruption of DNA-histone,

or histone-histone contacts, in order to decondense chromatin; and the second is to recruit a series of nonhistone cofactors. These proteins carry with them enzymatic activities, that are capable of further modifying chromatin and its structure (Kouzarides, 2007). Alterations in chromatin structure have been linked to functional outputs including cellular processes such as transcriptional regulation, DNA damage repair, and replication. Some of the extensively studied modifications will subsequently be discussed in detail.



**Figure 1.3. Overview of the different classes of modification identified on histones and their associated functional outputs.** Figure adapted from Kouzarides, 2007.

### 1.4.1. Acetylation

Acetylation of histone lysine residues is one of the most thoroughly investigated covalent modifications influencing gene expression in eukaryotic cells. Acetylation is a reversible modification where lysines undergo an acetylation-deacetylation switch governed by different physiological conditions. The transition between these modifications is controlled through the activities of two sets of enzymes named the lysine/histone acetyltransferases (KATs/HATs) and the lysine/histone deacetyltransferases (KDACs/HDACs). KATs function by catalysing the transfer of an acetyl group from acetyl-CoA molecules to the  $\epsilon$ -amino group of lysine residues within the N-terminal tails of histones (Berndsen et al., 2007).

Histone acetylation occurs throughout the whole eukaryotic genome and is generally associated with the activation of transcriptional events. Acetylation potentially neutralises electrostatic interactions between the histone lysines and phosphate groups of DNA, thereby preventing the compaction and folding of higher order chromatin structures, reducing internucleosomal contacts and promoting accessibility for the transcriptional machinery. Furthermore, lysine acetylation, similar to other PTMs, is significant as a direct signal and binding platform for trans-acting factors (Verdone et al., 2006). Antagonistically, KDAC activity and deacetylation cause repression at specific genes. The effect of deacetylation is partly mediated by a direct compaction of chromatin structure and also in promoting the binding of proteins, generally repressors, which recognise unacetylated histones. For example, heterochromatic spreading requires the recruitment of the Sir silencing complex by the active deacetylation of H4K16. This, in turn, promotes a condensed

structure and repression of gene expression (Shahbazian and Grunstein, 2007). Another important function of histone acetylation patterns is in assembly of chromatin. Newly synthesised histones are transiently acetylated, a factor which is important for their recognition by chaperone proteins and their correct deposition into nucleosomes, compacting replicated or newly transcribed DNA. After deposition, histones are rapidly deacetylated and later, reacetylated with new patterns to enable a variety of functions such as those stated above (Shahbazian and Grunstein, 2007).

As stated previously, PTMs function as an intricate language and it has been highly documented that the crosstalk between acetylation and other marks is fundamental in modulating chromatin-based transcriptional control and defining heritable epigenetic programs. This can be depicted by one of the first characterised examples of histone crosstalk between the histone H3 tail residues S10 phosphorylation and K14 acetylation. It was found that several KAT enzymes, in particular GCN5, displayed an increased acetyltransferase activity on H3 peptides at K14, when already bearing the S10 phosphorylation (Lo et al., 2000).

### 1.4.2. Phosphorylation

Phosphorylation is the reversible addition of a phosphate group to a serine or threonine residue within a protein substrate. Protein kinases catalyse the addition of phosphate groups and protein phosphatases remove those groups to modulate the balance between substrate phosphorylation and dephosphorylation within the cell. This modification affixes a bulky and negatively charged phosphate moiety to the hydroxyl side chain, thus increasing the electrostatic ion-pairing and hydrogen-

bonding capabilities of the serine, threonine and even tyrosine residues (Taverna et al., 2007). Alternatively, phosphorylation can induce protein interaction as it is read by specialised domains (Komander, 2009). Depending on the site of phosphorylation, this histone mark can function in the control of the cell-cycle, PTM-regulated signal transduction, DNA damage repair and transcriptional regulation. The phosphorylation of chromatin in cooperation with other histone marks has been shown in recent articles to have a direct effect on gene expression. Protein Kinase C $\beta$ 1 mediated phosphorylation of H3T6, for example, was shown to prevent LSD1 and JARID1B from demethylating H3K4 during androgen receptor gene activation (Metzger et al., 2010). A second illustration of crosstalk is that between H3S10ph and H3K9me3. At the onset of mitosis H3S10ph impedes an interaction between Heterochromatin Protein 1 (HP1) and H3K9me3, emitting HP1 from its binding site and antagonising transcriptional repression (Fischle et al., 2005).

### 1.4.3. Methylation

Methylation is defined as the transfer of a methyl group from S-adenosylmethionine (AdoMet) to the  $\epsilon$  amino group of a lysine residue or the guanidine nitrogen of an arginine residue (Bannister and Kouzarides, 2005). Methylation is generally more complex than acetylation as substrates can be mono-, di- or tri-methylated. These events are catalysed by three structurally defined types of lysine methyltransferase (KMT). The methylation of arginine residues is carried out by the Class I family of protein arginine *N*-methyltransferases (PRMT) and lysine methylation has been shown to be catalysed by the conserved SET domain family of

proteins. The stringent regulation of the activity and specificity of both sets of enzymes is, in part, carried out by different binding proteins and the complex in which it resides. The lysine methyltransferases will be discussed in more detail in section 1.12.

Histone methylation plays a fundamental role in a wide range of nuclear processes, including heterochromatin formation, X-chromosome inactivation and transcriptional regulation (Martin and Zhang, 2005). There is a wealth of research carried out on the subject of transcriptional control trying to decipher the methylation code. It has been found that whereas acetylation is predominantly associated with transcriptional activation, methylation can result in either activation or repression. The overall effect depends on the sequence-specific methylation site in histones and the methylation state of the arginine or lysine residues (Qian and Zhou, 2006). For example, three methylation sites are recognised to recruit activator proteins: H3K4, H3K36 and H3K79 and three are connected to transcriptional repression: H3K9, H3K27 and H4K20. Of course, as described in previous sections much remains to be discovered regarding how these signals work in combination to impact on gene regulation and other processes. Currently, twenty-four methylation sites have been identified in histones (Bannister and Kouzarides, 2005) however, the understanding of the mechanistic and functional consequences of this modification is limited. Nevertheless it can be speculated that the ultimate function of the methyl group is a reflection of the type of protein it has evolved to recruit.

Similar to acetylation, methylation is now known to be a reversible process. Work by Shi *et al.*, (2004) eliminated controversy in the field with the discovery of LSD1, a nuclear homolog of the amine oxidases. LSD1 acts to demethylate H3K4



and consequently repress transcription. However, in complex with the androgen receptor LSD1 demethylates H3K9 and activates transcription (Metzger et al., 2005). Recently, proteins harbouring the JmjC catalytic domain have also been found to demethylate lysine residues but via a distinct catalytic reaction to LSD1 (Kouzarides, 2007).

As of yet, arginine demethylases have not been identified however, the presence of lysine demethylases suggest it is likely arginine demethylases do exist. The lysine demethylase LSD1 acts via an amine oxidase reaction, a reaction predicted to be compatible with arginine demethylation also. Although not established, members within this enzyme family may be able to demethylate arginine residues (Bedford and Richard, 2005). It has also been shown that adjacent, pre-existing PTMs close to a site of methylation can mask the recognition motif of protein methyltransferases (Bedford and Richard, 2005). In addition to the above mechanisms, an alternative pathway for the reversal of arginine demethylation has been proposed. This process is termed deimination and involves the conversion of unmodified or mono-methyl arginine residues into citrulline at specific sites on the tail of histone H3 and H4 (Bannister and Kouzarides, 2005). Deimination is mediated by the enzyme peptidyl arginine deiminase 4 (PADI4) and prevents arginine methylation by CARM1. This modification represents a novel mechanism for antagonising the transcriptional induction mediated by arginine methylation (Cuthbert et al., 2004).

### 1.4.4. Ubiquitination

Protein ubiquitination involves the modification of lysine residues with the small protein ubiquitin, a polypeptide composed of 76 amino acids. Ubiquitin is covalently attached to lysine residues via a three-step enzymatic series. In an ATP-dependent reaction, ubiquitin activating enzymes (E1) append ubiquitin to an E2 conjugating enzyme which acts in ubiquitin transfer. Finally E3 ligases function as adaptors by binding both the substrate and E2-ubiquitin substrate; in this manner they function in the recognition of substrates and the facilitation of isopeptide bond formation between the substrate and ubiquitin (Komander, 2009). Within its own primary sequence ubiquitin contains seven lysines, rendering it also a substrate for ubiquitination. This gives rise to various combinations of mono- and poly-ubiquitination that can occur on a single substrate. Poly-ubiquitination adds greater complexity to the 'code' thereby increasing its functional versatility. Protein ubiquitination is a reversible modification, in which the removal and editing of ubiquitin chains is carried out by specialised proteases named deubiquitinating enzymes (Sowa et al., 2009). Initially this ubiquitin tag was found to serve as a signal for protein turnover as an ubiquitinated protein would undergo rapid degradation mediated by the proteasome. However, ubiquitination has since been shown to function in several other processes quite distinct from the regulation of protein half-life. Such processes include transcriptional control, DNA repair (Kouzarides, 2007), the DNA-damage response, endocytosis, cell signalling and trafficking and the mediation of protein-protein interactions (Komander, 2009).

### 1.4.5. SUMOylation

Small ubiquitin-related modifier (SUMO) is an 11kDa protein that can be coupled to other proteins to control their function. This modification has been shown to take place on all four histones, and specific sites have been identified on H2A, H2B and H4 (Nathan et al., 2006). SUMO resembles ubiquitin structurally but not functionally. Whereas ubiquitin tags proteins for degradation, SUMO has been shown to influence the targets function by affecting its intracellular localisation, interactions, stability and activity (Meulmeester and Melchior, 2008). The SUMOylation process has been implicated in the regulation of an array of fundamental cellular processes; including cell cycle regulation, gene transcription and cellular differentiation, and is essential for the well-being and survival of most organisms. Mammals express at least four SUMO proteins SUMO 1,2,3 and 4, which can be distinguished by their different target proteins and consequent differential functions (Meulmeester and Melchior, 2008). The reversible SUMOylation process begins with the activation of mature SUMO by the E1 activating enzyme Aos1-Uba2. SUMO is subsequently transferred to the E2 conjugating enzyme Ubc9. Ubc9 can make contact with SUMO interacting motifs (SIM) within target proteins and, facilitated by E3 ligases, SUMO finally forms an isopeptide bond with the amino group of a lysine residue in its target protein. Several E3 ligases have been identified, however, interestingly the PHD domain of the KAP1 corepressor has been shown to possess intramolecular E3 ligase activity. This activity is required for Ubc9 binding and the succeeding direction of SUMOylation to site-specific lysine residues within its adjacent bromodomain. SUMOylation is required for KAP1-mediated gene silencing (Ivanov et al., 2007).

Dynamic interplay has been shown between SUMOylation, acetylation and ubiquitylation which occur on the same lysine residue. SUMO therefore antagonises and acts as a potential block to activating modifications, coupling the modification with transcriptional repression (Nathan et al., 2006). SUMOylation can be reversed by SUMO-specific proteases, which cleave the isopeptide bond between SUMO and its target protein.

### 1.4.6. ADP Ribosylation

ADP ribosylation involves the transfer of the ADP-ribose moiety of ADP-ribosyl nicotinamide (NAD<sup>+</sup>) precursor to an arginine/ glutamic acid residue on an acceptor protein, also releasing the bi-product nicotinamide. This modification is reversible and exists in mono- and poly- forms. The enzymes that mediate it are the Mono-ADP-ribosyltransferases (MARTs) or the Poly-ADP-ribose polymerases (PARPs) respectively. In addition to these two classes of enzymes, the Sir family of NAD<sup>+</sup> dependent histone deacetylases have been shown to possess low levels of MART activity, so possibly represent a novel family of intracellular MARTs (Hassa et al., 2006). There are many reports of ADP-ribosylation of histones, however this modification remains ill-defined with respect to function. It has been linked to many fundamental physiological processes including signalling, transcription, DNA repair and apoptosis, whether these functional outputs result from MART/PARP enzymatic activity or the actual histone modification itself remains elusive. In terms of chromatin structure, it has been shown that poly-ADP ribosylation results in a relaxation at the site of modification, leading to an opening for factors involved in

DNA-mediated processes (Hassa et al., 2006). Evidence that catalytic activity is involved in such processes is prevalent for the role of PARP-1 in DNA repair. The presence of double strand breaks leads to an induction of 10-500 fold in the activity of PARP-1, which in turn leads to local changes in chromatin architecture (Ju et al., 2006).

### 1.4.7. Biotinylation

Evidence exists that lysine residues within histones are also modified by the covalent attachment of the vitamin biotin, catalysed by biotinidase and holocarboxylase synthetase (HCS). Eleven biotinylation sites have been identified in histones H2A, H3 and H4 and linked to biological functions such as cell proliferation, gene silencing and the cellular response to DNA damage. Specifically, H4K12 biotinylation is enriched in heterochromatic regions and plays a role in gene repression (Hassan and Zemleni, 2008). Observations have also shown that crosstalk exists between the biotinylation of histones and other modifications such as methylation and phosphorylation. For example, the dimethylation of R2, R8 and R17 lead to an increase of biotinidase-mediated biotinylation of K4, K9 and K18 (Kobza et al., 2005). Biotinylation is a reversible process; however the enzymes or mechanisms required for debiotinylation remain uncertain. Yet, the abundance of biotinylated histones at specific genomic loci does depend on the exogenous biotin supply and availability (Hassan and Zemleni, 2008). Recent doubt has been cast over this modification with regards to its *in vivo* occurrence. Healy and colleagues have accumulated data suggesting biotin is absent in native histones, suggesting that

the regulatory function of biotin on gene repression must be through alternate mechanisms (Healy et al., 2009).

### 1.4.8. Proline Isomerisation

Peptidyl proline exists in either a *cis* or *trans* conformation and conformational changes between the two states results in severe distortion of the polypeptide backbone. Proline isomerisation is a non-covalent modification that can occur both in histone and non-histone proteins, acting as a regulatory switch in signalling pathways. Isomerisation occurs spontaneously, however enzymes have evolved to accelerate the interconversion of the proline isomers. The enzyme, FPR4 a member of the FK506 binding proteins, has been identified in budding yeast and can bind the amino terminal tails of H3 and H4, specifically proline residues P30/38 of histone H3. Isomerisation of P38 antagonises the methylation of H3K36 by SET2 providing evidence for the crosstalk between lysine methylation and proline isomerisation in the regulation of transcription (Nelson et al., 2006). The parvulin (Pin1) family of enzymes catalyse the isomerisation of peptidyl-prolyl bonds in numerous transcription factors in a phosphorylation-dependent manner. In this case, isomerisation allows the control of the conformation, stability and activity of numerous proteins involved in transcriptional and cell cycle regulation (Shaw, 2007).

### 1.4.9. Histone Tail Clipping

Recently, a novel Histone H3 endopeptidase activity has been identified in *S. cerevisiae*. This led to the finding of a new type of histone modification, named ‘clipping’, involved in the regulation of gene expression. The enzyme cleaves histone H3 N-terminal tails after alanine residue 21 showing a bias for those tails carrying repressive signals; for instance those tails harbouring the activatory PTM H3K4me3 resist clipping (Santos-Rosa et al., 2009). Clipping has been implicated to function at promoter-bound nucleosomes therefore, suggesting a localised expulsion of repressive marks and proteins at promoters during the induction of expression. H3 tail clipping has also been shown to precede histone eviction, signifying the importance of this modification in marking which nucleosomes are to be displaced during transcriptional events (Santos-Rosa et al., 2009).

### 1.4.10. Chromatin Remodelling

As described previously, the above modifications serve as a code signalling for changes in the chromatin architecture and allowing the progress of fundamental nuclear processes. The dynamic properties of nucleosomes and consequent structural alterations are governed by the histone tails themselves and through the action of ATP-dependent chromatin-remodelling complexes. Alterations to histone proteins, especially the histone tails, were found to affect processes such as nucleosome sliding, histone dimer exchange and DNA wrapping (Ferreira et al., 2007). For example, truncation of the different histone tails yielded distinct observations for nucleosome mobility. Truncation of the H2A tail promotes nucleosome mobility,

whereas deletion of the H4 and H2B tails reduces it (Ferreira et al., 2007). In addition, despite having no effect on nucleosome mobility, the deletion of the H3 tail had the most prominent influence on nucleosome structure. Truncation of the H3 tail, with emphasis on residues comprising the  $\alpha$ -N helix, was found to alter the wrapping of nucleosomal DNA through a reduction in the stability of histone-DNA contacts within the nucleosome core. As a consequence, the removal of DNA from the octamer structure results in the destabilisation of H2A/H2B histone dimers within nucleosomes, facilitating access to the underlying genetic information (Ferreira et al., 2007). Chromatin remodelers utilise the energy of ATP-hydrolysis to shift, destabilise, expel or restructure nucleosomes in order to aid the creation of a dynamic chromatin environment (Clapier and Cairns, 2009). Distinctive targeting and tasks require specialisation and thus, remodelers are specific multi-protein complexes that can be defined, by their domain composition and their main functions, into four families; the Switching defective/Sucrose Nonfermenting (SWI/SNF), Imitation SWI (ISWI), Chromodomain, Helicase, DNA binding (CHD) and Inositol requiring 80 (INO80) remodelers. The families share a common catalytic subunit, the ATPase domain, which has been shown to use ATP-dependent DNA translocation to break histone-DNA contacts in order to redefine promoter architecture (Clapier and Cairns, 2009).

The ISWI family of remodelers aid in the organisation and assembly of chromatin, allowing the uniform spacing of nucleosomes at gene promoters. Generally, ISWI remodelling occurs at nucleosomes that lack acetylation, localising their activity to nucleosomes at transcriptionally inactive regions. ISWI complexes function by either disrupting nucleosome periodicity and generating disordered



nucleosome arrays, such as by the Nucleosome Remodelling Factor (NURF), or by creating regularly spaced, ordered nucleosome arrays, such as the function of the ATP-dependent Chromatin assembly and remodelling Factor (ACF), to uphold repressed or basal chromatin states (Cairns, 2009). In contrast the SWI/SNF complexes provide access to nucleosomal DNA and the exposure of DNA elements transiently on the surface. SWI/SNF family members can both slide and eject nucleosomes, correlating their functions with nucleosome disorganisation and promoter activation (Cairns, 2009). SWI/SNF complexes contain modular domains that recognise histone acetylation marks, increasing their targeting to and affinity for promoters undergoing activation. Furthermore, the mechanisms by which different remodelling enzymes act are sensitive to multiple aspects of nucleosome structure. For example, mutations to the histone H3  $\alpha$ N region selectively affected the rate of nucleosome repositioning and the products generated following ATP-dependent remodelling by Chromatin Structure Remodelling (RSC) complex, compared to a lesser effect with the Chd1 and SWI/SNF remodelers (Somers and Owen-Hughes, 2009).

Both these examples highlight the importance of PTM recognition by non-catalytic subunits of the remodeler complexes. These reader domains of the histone code are involved in the regulation of catalytic ATPase activity and the remodelling reactions and products by stabilising contacts with the nucleosome core surface, enhancing nucleosome sliding and dictating the directionality of nucleosome movement (Clapier and Cairns, 2009).

### 1.5. Lysine Acetyltransferase Enzymes

The importance of histone acetylation by HATs has long been recognised in chromatin biology; especially with respect to its correlation with transcriptional regulation, whereby hyperacetylation of histones is prominent in euchromatin. Dozens of proteins have been discovered which possess intrinsic acetyltransferase activity and of late, it has emerged that the activity of some of these enzymes is directed towards non-histone lysine residues. Studies on a global scale, including high-resolution mass spectrometry and the genome wide ChIP-Sequencing mapping of chromatin, have demonstrated how abundant protein acetylation is and how broad the regulatory effect is upon cellular processes (Choudhary et al., 2009). Due to this a new field of research is underway, exploring the entire set of acetylated proteins in a cell or otherwise referred to as the ‘acetylome’ (Smith and Workman, 2009). Such substrate diversity and rapid discovery of new chromatin modifying enzymes has founded a new coherent nomenclature system, where the more generic term of lysine acetyltransferases (KATs) is now used (Allis et al., 2007). KATs are a diverse superfamily of enzymes evolutionary conserved from yeast to humans and are characterised into different groups based upon sequence comparisons. Three main groups including the Gcn5-related N-acetyltransferase (GNAT), p300/CBP and MYST families are among those most intensely studied because of their broad conservation, varied functions and links with disease. These proteins function as transcriptional coactivators and in a diverse array of other nuclear and cytoplasmic processes.

KATs catalyse the addition of an acetyl moiety from an acetyl-CoA molecule to the  $\epsilon$ -amino group of a protein lysine residue. Enzyme acetyltransferase activity is regulated in a variety of ways including post-translational modification, degradation or altering the subcellular localisation of the KAT (Yang, 2004). In addition, most KATs physiologically exist as part of a multisubunit complex, in which different combinations of subunits contribute to the distinctive characteristics of each KAT complex (Lee and Workman, 2007). KAT complexes are generally more active than their counterpart enzymes and demonstrate distinctive substrate targeting, implicating the non-catalytic proteins in the regulation of activity and substrate specificity (Yang, 2004).

The specificity of KAT action still remains an elusive topic despite intensive characterisation of the different complexes. Biochemical and genetic analyses advocate that KATs exert a high degree of specificity towards target acetylation and in the processes they regulate. However, with the advancement of genome-wide mapping techniques such as ChIP combined with microarray (ChIP-chip) or with high-throughput sequencing (ChIP-Seq), it has been shown that KATs co-occur simultaneously to a similar set of loci in the genome, suggesting a low degree of specificity in their function (Anamika et al., 2010). Recently, a model has been proposed combining all data sets, in which KATs have a dual mode of action. This consists of an 'initiation' stage of transcriptional activation which is defined by the specific recruitment of one particular KAT to a given gene promoter. The sequential transition into 'maintenance' phase represents the stabilisation of an active transcription stage where recruitment of multiple KATs bind in a less-specific manner to the gene promoter (Anamika et al., 2010). The development of novel

ChIP-grade antibodies against KATs and complex-specific subunits will allow further investigation into the functional role and specificity of acetyltransferases *in vivo*.

### 1.5.1. MYST family

The MYST (MOZ, Ybf2/Sas3, Sas2, TIP60) family was originally named for its four founding members in yeast and mammals and at present consists of five human KATs: TIP60 (KAT5), MOZ (KAT6A), MORF (KAT6B), HBO1 (KAT7) and MOF (KAT8). This family of KATs are highly conserved in higher eukaryotes and in addition, orthologues have been identified in many lower eukaryotes. Sequence and structural comparisons depict their defining MYST catalytic domain which is composed of a zinc finger and an acetyl-CoA binding motif.

Obtaining consistent details of the catalytic mechanism for the MYST family of KATs has proven problematic. Different groups have carried out meticulous kinetic and structural analyses to unravel the mode of catalysis; however, discrepancies lie amid their individual findings. Essential Sas2-related acetyltransferase 1 (Esa1), the yeast homologue of TIP60, is used as the prototypical enzyme for understanding the structure and mechanism of the MYST family (Berndsen et al., 2007). Originally, based primarily on the structural similarity to the catalytic core of the GCN5 / p300-CBP associated factor (PCAF) subfamily, initial structural determination of Esa1 suggested a sequential mechanism of acetyl transfer (Yan et al., 2002). Subsequently, after additional structural and functional studies, Yan and colleagues re-assessed the mechanism, and put forward the new double-displacement model (ping-pong catalysis). This proceeds through the transfer of the

acetyl group of the acetyl-CoA cofactor to a cysteine shown to be strictly conserved among MYST family members. The acetyl group is then transferred from this acetyl-cysteine enzyme intermediate to a histone lysine substrate (Yan et al., 2002). The  $\epsilon$ -amino group ( $\text{N}\epsilon\text{H}_2$ ) of the target lysine is known to be deprotonated, prior to acetyl transfer by a glutamate residue conserved among all KATs, despite their differing catalytic mechanisms. In recent years, this mechanism was challenged by Berndsen and colleagues, whose data represents results yielded from the minimal core complex piccolo NuA4, and supports the original ternary complex mechanism involving a direct attack of a deprotonated N- $\epsilon$ -lysine on the bound acetyl-CoA (Berndsen et al., 2007). Mutational analysis is casting doubt on the importance of the conserved cysteine residue for MYST acetyltransferase activity; however currently there is still active controversy in the field, highlighting the need for further research. Such understanding into the mode of action of KAT enzymes will be crucial to the design of inhibitors. Besides the MYST domain other common structural features of the MYST proteins include chromodomains, plant homeodomain-linked (PHD) zinc fingers and regions defined by a high quantity of a particular amino acid residue (Figure 1.4).

MYST family enzymes conduct a considerable proportion of all nuclear acetylation and have a well-established role in histone acetylation. MYST-dependent regulation of chromatin impacts processes including gene regulation, DNA repair, development, cell cycle and stem cell homeostasis. However, recent studies show a broader substrate spectrum exists for MYST KATs and novel non-chromatin targets under their regulation are continually being discovered (Sapountzi and Cote, 2010). Of the five human KATs, TIP60 has the most known non-histone targets. With

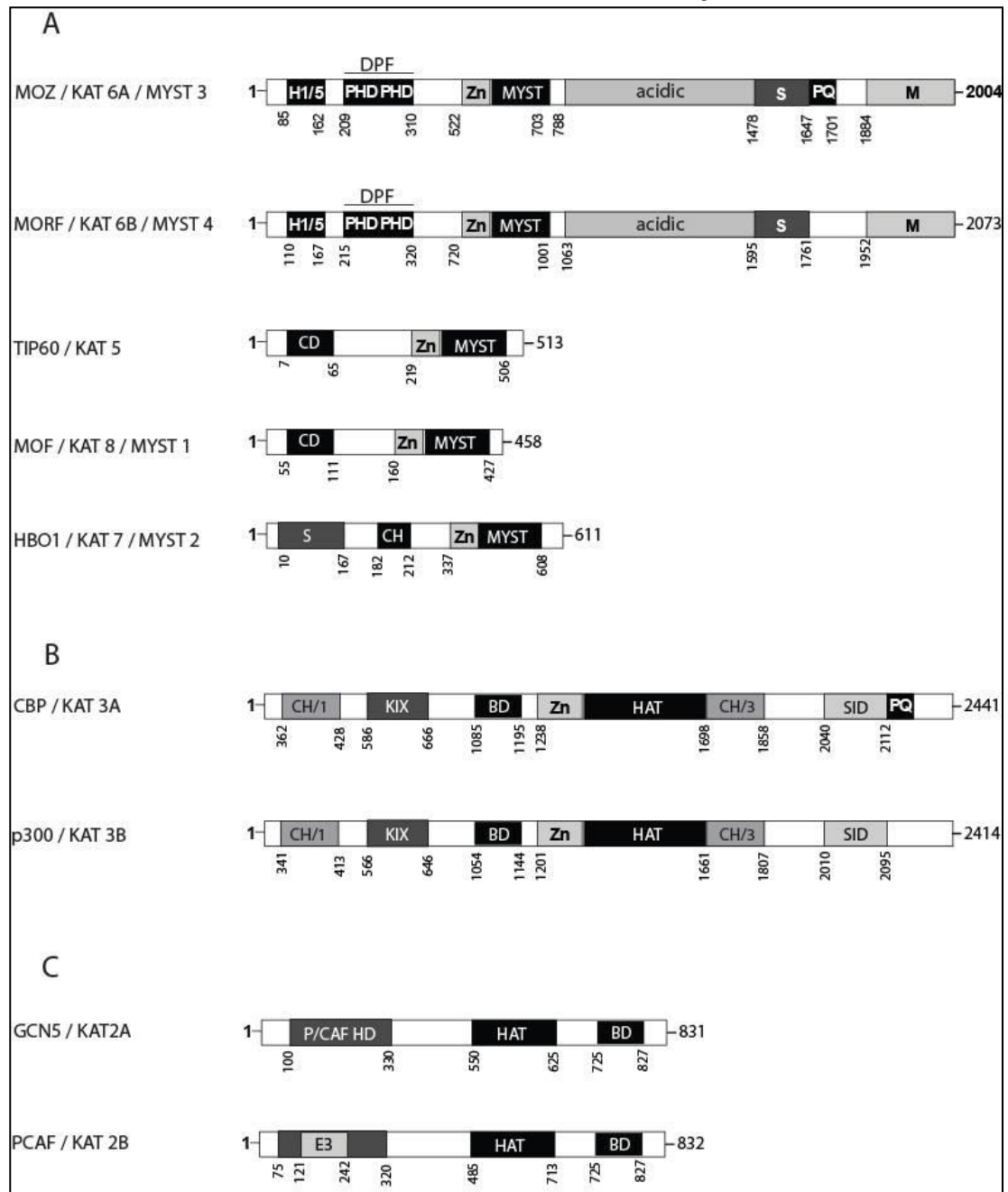
respect to chromatin-based targets the NuA4 complex, containing TIP60, is a key regulator of the transcription of genes involved in cell cycle control and embryonic stem cell identity, and also modifies chromatin surrounding DNA double strand breaks to aid repair (Squatrino et al., 2006). The majority of non-histone proteins acetylated by TIP60 are transcription factors, most of which are directly linked to transcriptional control. Acetylation by TIP60 exerts its effects through the alteration of protein substrate stability or activity (Sapountzi and Cote, 2010). TIP60 acts by the acetylation and activation of components of the DNA repair pathway, including p53 and ATM, modulating the decision between cell-cycle arrest and apoptosis upon p53 activation (Sykes et al., 2006). In contrast, the stability of the oncoprotein c-Myc is regulated by TIP60 acetylation, where a dramatic decrease in the rate of protein degradation is seen upon acetylation (Patel et al., 2004).

### 1.5.2. CBP/p300 family

The nuclear CREB-Binding Protein (CBP) was first identified as an interacting partner of the activated phosphorylated form of the CREB (cAMP response element binding protein) transcription factor (Chrivia et al., 1993). Its paralogue p300 was alternatively isolated as a cellular factor bound by the adenovirus E1A protein (Eckner et al., 1994). Both proteins are ubiquitously expressed in all higher eukaryotes, including flies, worms and plants, but not in lower eukaryotes such as yeast (Yuan and Giordano, 2002). CBP and p300 are similar proteins which exhibit several regions of close homology within their amino acid sequences (Figure

1.4). These include the cysteine-histidine-rich regions (CH1, -2 and -3), bromodomain, and the acetyltransferase domain.

CBP and p300 have been implicated in a host of cellular processes; both chromatin related such as transcription, DNA replication and repair and those not directly linked to chromatin, such as the regulation of p53 turnover in cycling cells (Grossman et al., 1998) and nuclear import (Ryan et al., 2006). Two important aspects, related to their modular structure, render CBP and p300 as effective global coactivators. By simultaneously interacting with the basal transcription machinery and other specific transcription factors, both proteins function as physical bridges which stabilise the transcription complex (Kalkhoven, 2004). Secondly, the ability of CBP/p300 to acetylate promoter proximal nucleosomal histones, results in the decompaction of local chromatin structure, allowing the access of other essential regulators (Kundu et al., 2000).



**Figure 1.4. Modular structure of Lysine AcetylTransferases (KATs).** Domain organisation of lysine acetyltransferase enzyme families (A) MYST (B) CBP/p300 (C) GCN5/ PCAF from humans, indicating highly conserved regions and functional domains. Domains are labelled as follows: H1/5, histones H1- and H5-like domain; PHD, plant homeodomain; DPF, double PHD finger; Zn, C2HC zinc finger; MYST, (MOZ, Ybf2/Sas3, Sas2, TIP60) homologous acetyltransferase region; acidic, glutamate/aspartate-rich region; S/M, serine/ methionine-rich region; PQ, proline/glutamine-stretch; CD, chromodomain; CH, cysteine/histidine rich motif; KIX, CREB binding region; BD, bromodomain; SID, steroid receptor coactivator 1 interaction domain; P/CAF HD, P/CAF homology domain; E3, E3 ligase domain.



The p300/CBP HAT domain is comprised of two important functional entities; the coenzyme A (CoA) binding and catalytic site, and zinc finger (Figure 1.4). Recently, ambiguity concerning the distinct catalytic mechanism of CBP/p300-mediated acetylation has been resolved, in part, by the production and analyses of high resolution crystal structures of semi-synthetic p300 in complex with its inhibitor Lys-CoA. When compared to other KATs, Liu and colleagues identified structural conservation within the central core region associated with CoA binding. However, several novel features separate these two coactivators from the other KAT families, and begin to provide reasoning for their characteristic broad substrate specificity. Here structural, kinetic and mutagenesis data proposes that two substrate binding pockets, connected via a narrow electronegative groove exist; one responsible for CoA binding and a second highly electronegative pocket required for protein substrate binding. It is suggested that catalysis occurs through the Theorell-Chance mechanism, where substrates weakly associate with the p300 surface, allowing the lysyl residue to advance through the groove and react with the acetyl group (Liu et al., 2008). It has been shown that regulation of the p300 HAT domain occurs via an activation loop. This loop, rich in basic residues, is thought to associate with the secondary binding pocket until extensive autoacetylation occurs on constituent poly-lysine residues (Liu et al., 2008). Autoacetylation induces specific structural changes in the acetyltransferase domain (Arif et al., 2007). These conformational changes, accordingly act as a master switch to expose the substrate binding site and enhance the catalytic activity of p300 (Thompson et al., 2004).

### 1.5.3. GNAT family

Homologues of GCN5 have been identified in a host of different organisms including both higher and lower eukaryotes. In mammals, two related GCN5 acetyltransferase subclasses were described: GCN5 and PCAF (Verdone et al., 2006). Both KATs are considered as global coactivators, and are indispensable for normal cellular function. Vertebrate GCN5 is approximately 73% identical to the vertebrate PCAF protein, with most identity focused into three highly homologous domain structures (figure 1.4). The N-terminal half contains the PCAF homology domain, and the C-terminal consists of the bromo- and acetyltransferase domains.

GCN5 and PCAF have been implicated in a host of processes including transcriptional activation, cell-cycle, differentiation and apoptosis. Further, PCAF acts as a co-activator in myogenesis (Puri et al., 1997), nuclear-receptor-mediated activation (Blanco et al., 1998) and growth-factor-signalled activation (Xu et al., 1998). Numerous substrates are known for GCN5 and PCAF, out of which their major targets are histones and nucleosomes. These enzymes exist as part of multisubunit complexes, such as the GCN5 complex containing  $\gamma$ SAGA, hTFTC, hSTAGA and the PCAF complex containing hPCAF. Residence within the protein complexes has been shown to alter their nucleosomal substrate specificity; for instance when incorporated into a complex, the acetylation target of GCN5 changes from H3K14 to H3K9, K18, K23, and to a lesser extent H2B and H4 (Guelman et al., 2006). As well as possessing locus-specific coactivator functions GCN5 and PCAF are able to acetylate non-histone substrates including p53 (Poux and Marmorstein, 2003), TFIIE/F (Imhof et al., 1997) and MyoD (Dilworth et al., 2004).

Recently, it was shown that along with acetyltransferase activity, PCAF possesses intrinsic ubiquitin E3 ligase potential within its PCAF homology domain. Such activity was demonstrated to control the stability of the oncoprotein Hdm2, a protein that promotes p53 degradation, and thus plays a role in regulating cellular p53 levels (Linares et al., 2007). Whether GCN5 also harbours this activity is yet to be determined.

The acetyltransferase domain of various GNAT family members has been extensively studied. This domain can be divided into two distinct regions; the structurally conserved central protein core associated with AcCoA binding, and a pronounced cleft flanked by N- and C-terminal regions. These N- and C-terminal segments show structural divergence from other family members, and are implicated for histone substrate binding and catalysis (Trievel et al., 1999). The catalytic mechanism of the GNAT superfamily was one of those first to be documented. Results from detailed bi-substrate kinetic analysis and product inhibition studies supported a sequential ternary complex (ordered Bi-Bi) kinetic mechanism (Tanner et al., 2000a; Tanner et al., 2000b). Acetyl-CoA binds first with a high affinity for the free form of the enzyme, followed successively by the H3 histone substrate. Within the ternary complex, the  $\epsilon$ -amino group of the lysine substrate directly attacks the carbonyl carbon of AcCoA, transferring the acetyl group to the acceptor peptide. Preceding catalysis, a conserved glutamate residue (hPCAF-E570, yGCN5-E173) acts as a general base catalyst, deprotonating the  $\epsilon$ -amino group and facilitating nucleophilic attack. The resulting products are then released sequentially, with the CoA following the ejection of the acetylated H3 product. The regulation of PCAF function proceeds through its own post-translation modification. Human PCAF is

acetylated, *in vivo*, by itself or p300 through intra- or intermolecular events, resulting in the augmentation of its acetyltransferase activity (Santos-Rosa et al., 2003).

The diverse superfamily of lysine acetyltransferases executes an acetylation spectrum that is essential for a diverse range of cellular processes. Therefore, disruption of such enzymes and their consequent acetylation patterns may lead to the development of cancer and other diseases.

### 1.6. Structure and Function of MOZ / MYST3

The monocytic leukaemia zinc finger protein (MOZ), also called MYST3 or KAT6A, is a histone acetyltransferase belonging to the MYST family of KATs. MOZ was first identified in recurrent chromosomal translocations which involve the fusion of *MOZ* to the gene encoding *CBP*; an event associated with an aggressive form of acute myeloid leukaemia (AML) (Borrow et al., 1996).

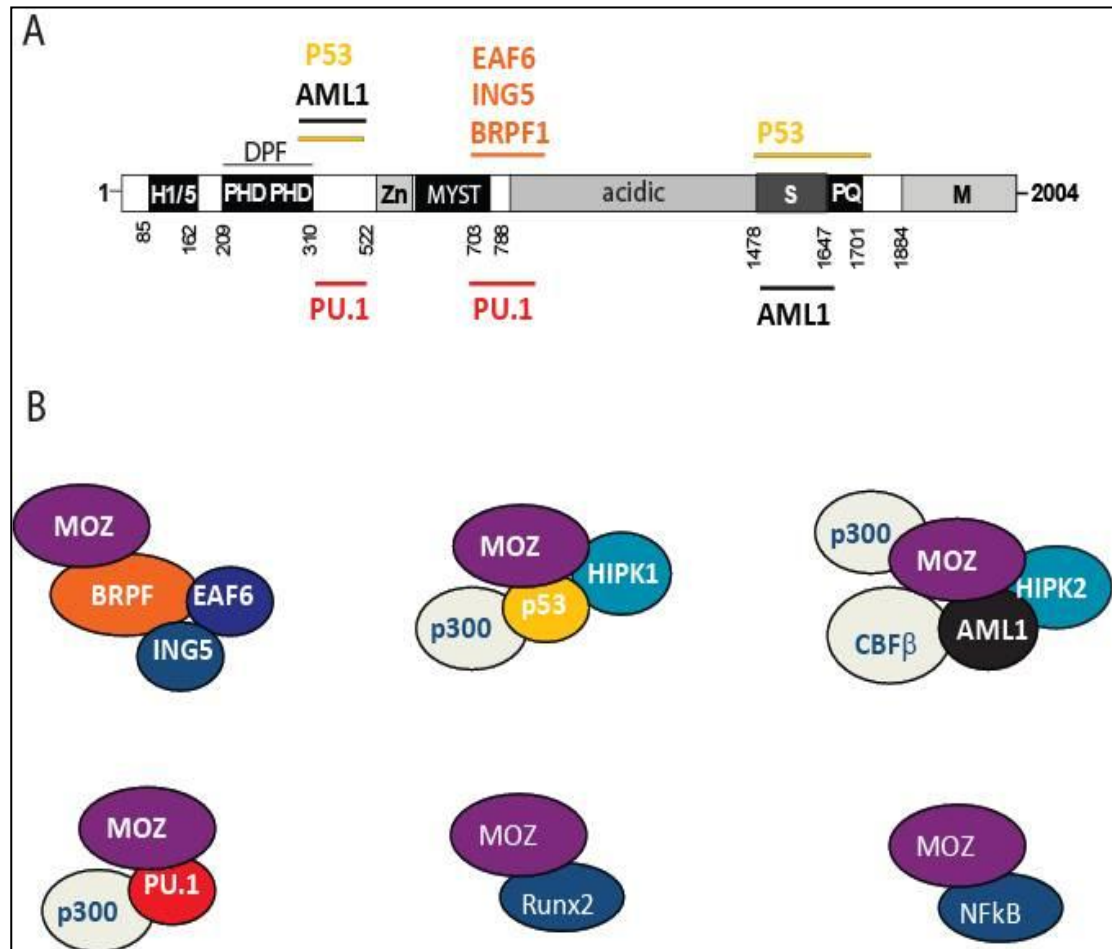
#### 1.6.1. MOZ is a modular protein

The modular structure of MOZ consists of several discrete domains (Figure 1.5(A)). The central MYST histone acetyltransferase domain, which it shares with the other four mammalian MYST family members, is thought to perform the catalytic function of the enzyme. Recently, novel results suggest that enzymatic catalysis and DNA targeting activities are both contained in the MYST regulatory domain; a feature not conserved among other histone acetyltransferases. Probing by Holbert and colleagues into the MOZ MYST domain, by means of structural and biochemical studies, revealed a conserved central core associated with the binding of acetyl-CoA

and catalysis. This core is flanked by evolutionary divergent N- and C-terminal regions harbouring a TFIIIA-type zinc finger and helix-turn-helix DNA-binding motifs respectively. MOZ was shown to bind to DNA through these motifs and furthermore, kinetic data indicated a distinct surface for DNA binding separate from that of catalysis; implying the two MYST domain properties are not mutually exclusive (Holbert et al., 2007).

Amino-terminal of this are a H1/5 like domain (also named N-terminal part of ENOK- NEMM) and tandem plant homeodomain zinc fingers (PHD). The H1/5 domain shows a weak sequence homology to motifs seen in the linker histones H1 and H5, although its function still remains somewhat ambiguous. Yet, through deletion analyses, the Kitabayshi group has implicated the H1/5 domain in nuclear localisation and emphasises its essential function in transcriptional coactivation (Yoshida and Kitabayashi, 2008). Since this domain has also been associated with a function in nucleosome binding (Ramakrishnan et al., 1993), it is thought that MOZ may be able to directly alter chromatin compaction by interaction with the nucleosome. The functional importance of the double PHD fingers (DPF) is unclear, yet it is speculated that they recognise methyl-lysine-containing motifs of histones (Kouzarides, 2007). More recently however, the characterisation of a tandem PHD domain in the recognition of acetylated histone H3 shows the importance of two PHD fingers acting as one integrated functional unit (Zeng et al., 2010). The carboxy-terminal region of MOZ contains serine-, proline- and methionine-rich areas which have been reported to possess transcriptional activation properties *in vitro* (Champagne et al., 2001).

Attributable to the MYST domain, MOZ has been reported to possess intrinsic KAT activity towards all four core histones *in vitro* (Champagne et al., 2001; Holbert et al., 2007). MOZ can also acetylate histone H4 lysine residues 5, 8, 12 and 16 and histone H3 lysine 14 *in vitro* (Kitabayashi et al., 2001a). Separate studies have also highlighted histone H3K14 in HeLa cells (Doyon et al., 2006), histone H4K16 when comparing normal lymphocytes to leukaemia cells (Fraga et al., 2005), and histone H3K9 at *Hox* gene loci (Voss et al., 2009) as preferential targets of MOZ acetylation. Such studies reveal a broad range of MOZ-specific acetylation targets, causing controversy as to the actual identity of lysine residues MOZ may show preference to acetylate. Further, functional data linking the HAT activity of MOZ at specific lysine residues to the regulation of transcription *in vivo* is limited. More recently, MOZ has been shown not to restrict its acetylation to only histone substrates but also to non-histone protein lysine residues. New acetylation targets include the complex component BRPF1 (Ullah et al., 2008) and the tumour suppressor p53 (Collins and Heery, unpublished data), however the regulation exerted by or the function of these PTMs is yet to be unveiled.



**Figure 1.5. Modular structure of MOZ and its interacting partners. (A)** Schematic representation of MOZ Domain organisation, indicating highly conserved regions and functional domains. Domains are labelled as follows: H1/5, histones H1- and H5-like domain; PHD, plant homeodomain; DPF, double PHD finger; Zn, C2HC zinc finger; MYST homologous acetyltransferase region; acidic, glutamate/aspartate-rich region; S/M, serine/ methionine-rich region; PQ, proline/glutamine-stretch. The mapped binding sites of published interacting partners are also detailed on the schematic. **(B)** MOZ functions as the catalytic subunit of a variety of multi-subunit complexes.

### 1.6.2. MOZ resides in multi-subunit complexes

MOZ resides as a catalytic subunit in several stable, nuclear multi-subunit complexes (Figure 1.5(B)). Kitabayashi *et al*, first isolated MOZ as part of the AML1-CBF $\beta$  transcription factor complex, whereby MOZ functions to stimulate AML1-mediated transcription (Kitabayashi et al., 2001a). Since then, a number of interacting partners have been identified and shown to function in complex with MOZ. These complexes will be discussed in subsequent sections related to their function.

The most recently discovered multi-subunit complex containing MOZ, also consists of the homolog of Esa1-associated factor 6 (EAF6) and Bromodomain- and PHD finger-containing (BRPF) proteins 1/2/3; BRPF are a family of proteins rich in modules commonly found in chromatin-associated factors including PHD fingers, bromodomain and the chromo/Tudor-related PWWP domain. The presence of such domains suggests that the BRPF protein is accountable for recruiting the MOZ complex to its site of action and consequently in chromatin regulation (Yang and Ullah, 2007). The inhibitor of growth 5 (ING5) member of the ING family of putative tumour suppressors is also co-purified as part of the MOZ complex *in vivo*. The ING family of tumour suppressors are known to be involved in many critical processes, however as they do not possess intrinsic enzymatic activity they are thought to function as adaptors, providing a powerful link between chromatin-modifying activities and their genomic sites. More specifically, ING subunits have been shown to be crucial for the KAT acetylation of chromatin substrates, where they are possibly



responsible for the initial binding to the chromatin through their PHD fingers (Doyon et al., 2006).

Further insight into the molecular architecture of the quartet MOZ KAT complex has recently been provided, revealing that the associated subunits regulate both the acetyltransferase and other functions of MOZ. BRPF1 has been shown to bind the MOZ MYST domain, which in turn acts as a scaffold for the association of ING5 and EAF6 subunits and the overall assembly of the complex (Ullah et al., 2008). At the functional level, complex formation significantly stimulated the acetyltransferase activity of the MOZ MYST domain and further influenced its preferred target. The native complexes differ in substrate specificity to their enzyme counterpart, primarily acetylating H3 and H4 free histones and nucleosomal histone H3 (Ullah et al., 2008). In addition, interaction with BRPF1 was also seen to augment the transcriptional activation potential of MOZ towards Runx2. Complex reconstitution studies are proving imperative to our understanding of the molecular mechanisms underlying MOZ function; where it is evident that under physiological conditions non-catalytic subunits act to regulate acetyltransferase and coactivator activities and refine substrate specificity.

### **1.6.3. MOZ functions as a transcriptional activator in hematopoietic processes**

MOZ interacts with many transcription factors in particular with hematopoietic specificity, such as Runx1 (AML1) (Kitabayashi et al., 2001a; Pelletier et al., 2002) and the ETS family factor PU.1 (Katsumoto et al., 2006); but also Runx2

(Pelletier et al., 2002), p53 (Rokudai et al., 2009) and NF- $\kappa$ B (Chan et al., 2007) functioning as a modest transcriptional coactivator (Figure 1.5). MOZ cooperatively activates transcription factor-dependent expression of genes (Collins et al., 2006), such as AML1-dependent transcription of myeloperoxidase (MPO) (Kitabayashi et al., 2001a); however the role of HAT activity in MOZ coactivator function remains unclear. As an example, the MYST domain is required for NF- $\kappa$ B-dependent transcription (Chan et al., 2007) but not for AML1-dependent transcription (Kitabayashi et al., 2001a), indicating that the role of HAT activity in gene regulation is dependent on the promoter context.

AML1 and PU.1 play fundamental roles in the establishment of definitive hematopoiesis, therefore as a coactivator of these transcription factors the role of MOZ in normal and leukaemic blood formation has been extensively studied. Gene targeting in mice demonstrated that MOZ-deficient mice are embryonic lethal at E15 with an undersized liver, the organ responsible for the production of hematopoietic cells during this developmental stage (Katsumoto et al., 2006). Flow cytometry was carried out on cells prepared from E14.5 fetal liver to observe the effect of the MOZ null mutation (MOZ<sup>-/-</sup>) on different populations of cells. Firstly, reductions were seen in the number of short-term hematopoietic stem cells (HSCs) and lineage-committed progenitors, implicating MOZ in the generation and maintenance of these early populations (Katsumoto et al., 2006). More specifically, in lineage-committed cells the number of B-lineage cells was severely reduced, maturation arrest was observed in erythroid lineage cells by an accumulation of erythroblasts and finally, an increased population of myeloid lineage cells was seen (Katsumoto et al., 2006). Increases in the committed myeloid cells suggest that differentiation is stalled at the

progenitor stage in the MOZ<sup>-/-</sup> fetal livers. Altogether, these observations highlight that MOZ plays a key role for appropriate development and differentiation of lymphoid and myeloid cells (Katsumoto et al., 2008).

An attribute of HSCs is their capability to reconstitute hematopoiesis of a lethally irradiated mouse. Competitive reconstitution assays were performed in order to establish whether MOZ<sup>-/-</sup> embryos possess functional HSCs, where these embryos were injected into irradiated normal recipient mice. Survival was therefore dependent on the ability of the donor cells to restore hematopoiesis. Their findings showed that no mice receiving cells from MOZ-deficient embryos survived past 28 days, indicating that MOZ is essential for sustaining the long-term repopulation of stem cells. MOZ therefore, regulates HSC self-renewal activity to reconstitute the hematopoietic system in recipients after transplantation (Katsumoto et al., 2006; Thomas et al., 2006).

As mentioned previously, the biological function of the well-established MOZ acetyltransferase activity is still ambiguous. Functional analyses investigating the role of MOZ in normal hematopoiesis was carried out in mice deficient of the entire MOZ protein. Defects observed in these studies are therefore likely to be a reflection of the cumulative effects of the loss of AML1 transcriptional coactivator function and HAT activity. Perez-Campo and colleagues used a mouse strain carrying a single amino acid mutation that inactivates HAT activity (HAT<sup>-/-</sup>), to explore the specific significance of MOZ-driven acetylation in the regulation of hematopoiesis (Perez-Campo et al., 2009). This study showed that mice exclusively lacking MOZ HAT activity exhibit a reduction in the number of HSCs and lineage-committed precursors, as well as a B-cell development defect. Additionally, hematopoietic precursors had a

profound deficit in their proliferative capacity, explaining the failure to maintain a normal number of precursors in HAT<sup>-/-</sup> cells (Perez-Campo et al., 2009). Taken together, this study demonstrates a crucial role for MOZ HAT activity in the proliferation and maintenance of hematopoietic precursors at all stages of development.

### 1.6.4. MOZ functions as a transcriptional activator in developmental processes

In addition to MOZ mRNA being detected in all hematopoietic progenitors, the gene is also extensively expressed throughout embryonic development and in adult tissues (Thomas et al., 2006). This implies that MOZ can regulate many biological processes both within hematopoiesis and in other cell types. In fact, MOZ has been implicated in developmental processes via the regulation of *Homeobox* gene expression, the gene clusters responsible for body segment identity during development. A zebrafish MOZ mutant phenotype was shown to exhibit developmental defects in the head region, where the mutation of MOZ results in the abolition of *Hox* gene expression specifically in skeletogenic cranial neural crest cells (CNC). This MOZ-dependent *Hox* gene expression in CNC determines the normal facial support skeleton in Zebrafish (Crump et al., 2006). In contrast, in MOZ-deficient mice, an extensive anterior homeotic transformation of the axial skeleton (neck and trunk region) and nervous system is observed (Voss et al., 2009). The severity of the developmental defect implicates MOZ as a regulator for a large number of HOX proteins. Results from Voss and colleagues confirmed that MOZ

deletion leads to a global decrease of *Hox* gene expression and further, effects the specification of the anterior expression boundary of *Hox* genes (Voss et al., 2009). ChIP using antibodies against specific histone PTMs, followed by quantitative genomic PCR revealed a reduction of H3K9 acetylation at the TSS of *Hox* loci in MOZ-deficient mice. Contradictory to previous *in vitro* studies, no effect on H3K14 acetylation levels was shown and H4K16 acetylation and H3K9 trimethylation were seen to increase in the MOZ<sup>-/-</sup> mutant compared to wild-type controls. MOZ is therefore required to maintain normal levels of H3K9 acetylation at *Hox* loci *in vivo* (Voss et al., 2009). Further novel findings disclose that H3K9 hypoacetylation, *Hox* gene repression and the anterior homeotic transformation caused by a lack of MOZ are all rescued by Retinoic acid (RA) signalling, a known potent activator of *Hox* gene expression. RA and MOZ are thus defined to act in parallel to define body segment identity (Voss et al., 2009).

The role of *Hox* gene expression in embryogenesis and development is well documented however, the homeobox genes have also been demonstrated to exert a critical function in normal hematopoiesis (van Oostveen et al., 1999). The abnormal expression of HOX transcription factors has been linked to both (Mixed Lineage Leukaemia) MLL-rearranged (Ayton and Cleary, 2003) and MOZ/CBP-associated leukemogenesis (Camos et al., 2006). A recent study explores the molecular crosstalk involved at the *Hox* gene promoters in order to regulate their expression in human cord blood CD34<sup>+</sup> cells (Paggetti et al., 2010). MOZ was found to colocalise and interact with the histone methyltransferase MLL and the adaptor protein essential for MLL-mediated H3K4 trimethylation, WDR5. Moreover, both modifying enzymes are recruited to multiple *Hox* promoters where they were found to synergistically

stimulate promoter activity (Paggetti et al., 2010). MOZ/MLL-targeted knockdown in multipotent CD34<sup>+</sup> cells was seen to alter the recruitment of the enzymes to *Hox* promoters, thereby decreasing the levels of H3K4me2/3 epigenetic modifications and consequently leading to *Hox* gene repression and an impaired ability to differentiate into myeloid progenitors (Paggetti et al., 2010). Dissecting the molecular mechanisms behind the regulation of MOZ-target genes is important for our understanding of MOZ function and further emphasises new ways by which MOZ/MLL crosstalk deregulation can lead to leukemogenic abnormalities in hematopoietic cells.

### **1.6.5. MOZ functions as a transcriptional activator in the DNA damage response**

MOZ interacts with p53, functioning as a coactivator for p53-mediated transactivation of specific target genes. The p53 protein is an essential regulator of the cellular response to different forms of stress, where it can trigger apoptosis to eradicate the damaged cell or initiate G<sub>1</sub> cell-cycle arrest to allow for damage repair. Kitabayashi's group investigated the role of the MOZ-p53 interaction in response to genotoxic stress (Rokudai et al., 2009). Flow cytometry experiments showed an impaired G<sub>1</sub> arrest of MOZ deficient cells following adriamycin (ADR)-induced DNA damage compared to wild-type, indicating that MOZ regulates this cell-cycle block. Further, an increase in the level of the p53-MOZ complex was observed in response to this genotoxic treatment. When examining the effect of MOZ<sup>-/-</sup> on the expression of different p53 target genes it was seen that only p21 expression was abolished. The p53-MOZ interaction is therefore required for the recruitment to and

activation of the p21-gene promoter in response to DNA damage and the consequent p21 expression leads to G<sub>1</sub> cell-cycle arrest (Rokudai et al., 2009). Interestingly, a recognised oncogenic mutation of p53, G279E, was seen to disrupt the p53-MOZ interaction but maintain the DNA binding activity of p53. This provides possible reasoning for the lack of ability of the p53 G279E mutant to activate transcription, where MOZ is required as its coactivator (Rokudai et al., 2009). The importance of p53/MOZ-mediated transcription is reinforced in this study, where its inhibition is involved in carcinogenesis.

The vital transcriptional coactivator function of MOZ is described above where MOZ-regulated genes are involved in processes including hematopoiesis, development, the DNA damage response (DDR) and carcinogenesis. Whereas our knowledge is constantly expanding, research into defining the relevance of all the MOZ modular domains and physiological complex components is required. Further, the refinement of both histone and non-histone lysine substrates and the mechanisms behind the control of MOZ-target genes is essential to establish a complete picture of MOZ function and its ultimate role in disease.

### **1.7. The Role of MOZ & Other KATs in Disease**

The activities of KAT and KDAC enzymes are vital for normal gene regulation and cellular homeostasis. Dysfunction of KATs has been linked to malignant transformation and many other diseases including cancer, diabetes, AIDS and neurodegeneration. More specifically, the MYST and CBP/p300 family members have been shown to play a causative role in leukaemogenesis. Acute myeloid

leukaemia is associated with frequent, reciprocal chromosomal translocations in hematopoietic progenitor cells. MOZ was first identified in recurrent chromosomal translocations which involve the fusion of *MOZ* to the gene encoding *CBP* (Borrow et al., 1996) and since then MOZ fusions incorporating the nuclear coactivators p300 (Kitabayashi et al., 2001b), TIF2 (transcriptional intermediary factor)(Carapeti et al., 1998) and NCOA3 (Esteyries et al., 2008) have been identified. AML that express MOZ-related translocations are generally monocytic leukaemias classified under the M4/M5 French-American-British subtypes (Yang and Ullah, 2007) and these MOZ fusions are found in approximately 6.5% of such AML subtypes (Esteyries et al., 2008).

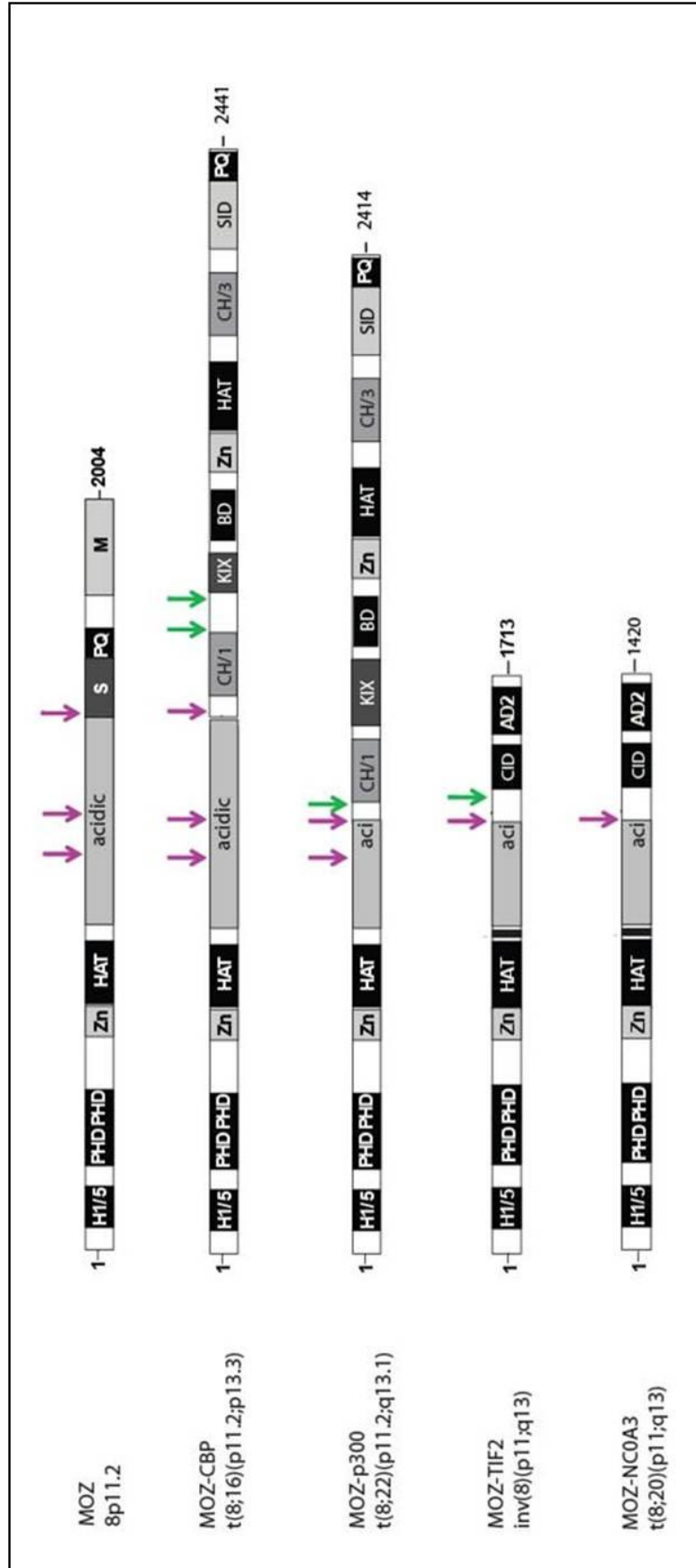
In all MOZ translocations identified to date, the breakpoint of MOZ is located within or around its acidic region (Figure 1.6). The N-terminal region of MOZ, containing the H1/5 like-, DPF-, MYST- and basic-domains is therefore retained and the C-terminus is replaced with one of the fusion partners listed above (Katsumoto et al., 2008). At the molecular level, the resulting creation of such chimeric products allows the catalytic subunits of one fusion partner to become mistargeted by the binding domains within the other. This can lead to a reduction of enzymatic activity at intended genomic loci, aberrant global and promoter-specific acetylation, sequestration of essential nuclear regulators (Avvakumov and Cote, 2007) or the instability of co-regulator complexes (Kindle et al., 2005), thus, perturbing gene expression programs necessary for normal myeloid differentiation.

MOZ-TIF2 and MOZ-CBP have been shown to exert their oncogenic effects by altering cofactor recruitment and chromatin modification at target promoters. Specifically, MOZ-TIF2 acts as a dominant inhibitor of the transcriptional activities



of CBP-dependent activators such as nuclear receptors and p53. This effect is caused by the direct binding of CBP to the C-terminal activation domain (AD1) of MOZ-TIF2 and as a consequence of this interaction, intracellular levels of CBP are strongly reduced and more specifically, CBP is depleted from PML bodies (Kindle et al., 2005). In addition, it has been shown that MOZ-TIF2 associates with the *RARβ2* promoter perturbing the ligand-dependent recruitment of CBP and resulting in the down-regulation of the *RARβ2* gene. In contrast, MOZ-TIF2 enhanced AML1/Runx1 reporter activation (Collins et al., 2006). The MOZ-CBP fusion protein shows similar transcriptional regulation functions at the *RARβ2* and AML1 promoters as compared to MOZ-TIF2 (Collins et al., 2006), but has also been implicated in the stimulation of NF-κB- dependent transcription (Chan et al., 2007). MOZ fusion proteins are therefore seen to differentially regulate the transcription of their targets, with transcription effects being possibly dependent on the promoter context.

The mechanism by which MOZ-TIF2 depletes intracellular levels of CBP and p300 is still ambiguous, however recently it has been shown to act via a proteasome-independent mechanism. Treatment of cells with a cell permeant calpain/cathepsin inhibitor (CPI-2) blocks the effect of MOZ-TIF2-mediated destruction of CBP/p300, restoring CBP expression to normal levels and rescuing MOZ-TIF2 mediated gene repression (Kindle et al., 2010). These results indicate a role for calpain and cathepsin proteases in the oncogenic action of MOZ-TIF2.



**Figure 1.6. Schematics of Leukaemia-associated MOZ Fusion Proteins.** MOZ contains several functional domains including the H1/5 (Histone1/5-like), PHD (plant homeodomain), Zn (zinc finger), HAT (histone acetyltransferase), acidic domains and S- (serine), PQ- (proline/glutamine), M- (methionine) rich domains. MOZ fuses with four known partners: CBP, p300, TIF2 and the recently identified NCOA3. MOZ fusion partners contain CH (cysteine/histidine rich), KIX (kinase-inducible), BD (bromodomain), SID (SRC1 interacting domain), CID (CBP interacting domain) and AD (activation domain). Chromosomal translocation breakpoints in the MOZ gene and its fusion partners are described as chromosome loci and mapped as pink arrows and green arrows respectively.

Evidence has shown that the binding of MOZ-TIF2 to the CBP cofactor also acts in extending the proliferative capacity of hematopoietic progenitors *in vitro* and induces AML *in vivo* (Kindle et al., 2005). When introduced into irradiated mice, MOZ-TIF2 is able to immortalise committed myeloid progenitor cells and cause AML (Huntly et al., 2004). Furthermore, it was shown that the C2HC nucleosome recognition motif of MOZ (Figure 1.6 (Zn)) is essential for the transforming properties of MOZ-TIF2, whereas the HAT activity of the MYST domain is dispensable (Deguchi et al., 2003). As well as interacting with CBP, it has most recently been shown that both MOZ-TIF2/CBP fusion proteins can interact with the PU.1 transcription factor to stimulate the expression of macrophage colony-stimulating factor receptor (CSF1R). Cells containing elevated levels of CSF1R showed potent leukaemia-initiating activity and ablation of such stem cells suppressed the progression of MOZ-TIF2 induced AML (Aikawa et al., 2010).

The role of KATs in cancer is complex, as, for example, it appears that the genes encoding CBP and p300 have both tumour-suppressor and tumour-promoter properties. A causative role for p300 and CBP in cancer has been implicated via many mechanisms. Firstly, both coactivators and in addition PCAF, are targeted by viral oncoproteins such as E1A. E1A binds and sequesters the KATs, inhibiting their acetyltransferase activities and preventing the binding of certain transcription factors. This represses transcriptional activation by transcription factors that require these KATs as coactivators (Frisch and Mymryk, 2002). Secondly, via the formation of chimeric fusion proteins, including those described above and also translocation with the mixed lineage leukaemia (MLL) gene, leads to the development of disease. The MLL-CBP/p300 fusion proteins are almost exclusively associated with

myeloproliferative diseases and have been proposed to lead to leukaemia by increasing histone acetylation of genomic regions targeted by MLL (Daser and Rabbitts, 2005).

Finally, a variety of p300 and CBP sequence alterations have been identified in malignancies. A loss of heterozygosity of p300 is frequently coupled with glioblastomas, colon, breast and ovarian cancers (Bryan et al., 2002). p300 mutations have also been described which predict a truncated protein present in epithelial and colorectal cancers. In addition, a somatic in-frame insertion of 6 amino acids into the p300 HAT domain was found in a primary breast cancer and missense alterations found in a primary colorectal cancer (Gayther et al., 2000). Studies using CBP knockout mice show an increased incidence of some cancers including leukaemia, but also defects in cell proliferation (Kung et al., 2000). Dysfunctional CBP has also been implicated in the developmental disorder Rubinstein-Taybi syndrome. Monoallelic mutation of the human CBP locus is sufficient to induce the on-set of this disease and patients with this syndrome exhibit an elevated risk of developing malignant tumours. Mutational analyses of patients reveal heterozygous mutations in the PHD finger and HAT domain of CBP, leading consequently to reduced acetyltransferase activity for its substrates and coactivator function for CREB (Kalkhoven et al., 2003). More recently, inactivating mutations of CBP and p300 have been linked to relapsed acute lymphoblastic leukaemia (ALL) (Mullighan et al., 2011) and multiple forms of B-cell non-Hodgkin's lymphoma (Pasqualucci et al., 2011). In both cases, deleterious substitutions or genomic deletions in the conserved residues of the HAT domain render it inactive, causing impaired histone acetylation and transcriptional regulation of CBP/p300 target genes.

With KATs playing a fundamental role in multiple cellular processes it is inevitable that their dysregulation will lead to the development of cancer and other diseases. Yet promisingly, our expanding understanding of KAT functional mechanisms is highlighting the curative value of these enzymes for the design of therapeutic interventions.

### 1.8. Epigenetic Modulation: Potential Anticancer Therapy

As described previously, there is a wealth of evidence indicating an aberrant acetylation status in cancer, whereby the complex interplay of KATs and KDACs is in some way perturbed. For this reason epigenetic drugs targeting KDAC chromatin modifying enzymes have been extensively studied and several are currently in clinical trials. Yet, it is important to appreciate that there are cancer cases involving the hyperacetylation of targets and the consequent use of KDACi in such situations could prove to be detrimental. At present, comparatively there is little information available on the anti-tumour effects of KAT inhibitors (KATi). This is most likely a result of i) KDAC inhibitors are still more efficient than KAT inhibitors; ii) information on the molecular basis of KAT inhibition is still vague; iii) the doses required for inhibition are high, causing difficulty in applying the inhibitors to cellular and biological systems (Manzo et al., 2009). However, with ever advancing knowledge of their structure, function and regulation KATs are emerging as promising pharmacological targets (Heery and Fischer, 2007).

The first reported KAT chemical inhibitors are bisubstrate analogues, in which a histone substrate peptide is covalently linked to a CoA motif at the lysine site

(figure 1.7A) (Zheng et al., 2008). Due to differential substrate specificities and catalytic mechanisms utilised by KATs, selectivity can be achieved through the alteration of the lysyl moiety in the conjugates. Lys-CoA was found to be a p300-specific inhibitor and H3-CoA-20 shows  $\mu\text{M}$  potency for PCAF (Lau et al., 2000). Despite their high potency for specific KATs these inhibitors exhibit low cell permeability and metabolic instability, which reduces their suitability for *in vivo* characterisation. To address this issue, the peptide CoA inhibitor was fused with a Tat peptide to yield a cell-permeable molecule and facilitate cell entry (Zheng et al., 2005). It remains to be seen if these constructs are pharmacologically useful, however these inhibitors have proven valuable in the elucidation of KAT biological function. KATs such as p300 have been co-crystallised with lysyl-CoA inhibitors and such experiments have provided mechanistic data and implicated p300 in novel regulatory functions (Liu et al., 2008). With respect to its therapeutic potential, this structural information could also inspire structure-based inhibitor design.

Enzymatic screens with natural products have yielded another category of small-molecule natural enzyme inhibitors. These include anacardic acid, curcumin and garcinol (figure 1.7B). Anacardic acid is a major component of cashew nutshell liquid and was identified as a non-specific p300/PCAF inhibitor (Balasubramanyam et al., 2003). However, again due to poor membrane permeability its *in vivo* effects could not be observed. Progressively, the scaffold of anacardic acid has been efficiently utilised by researchers to synthesise several KATi (Selvi and Kundu, 2009).

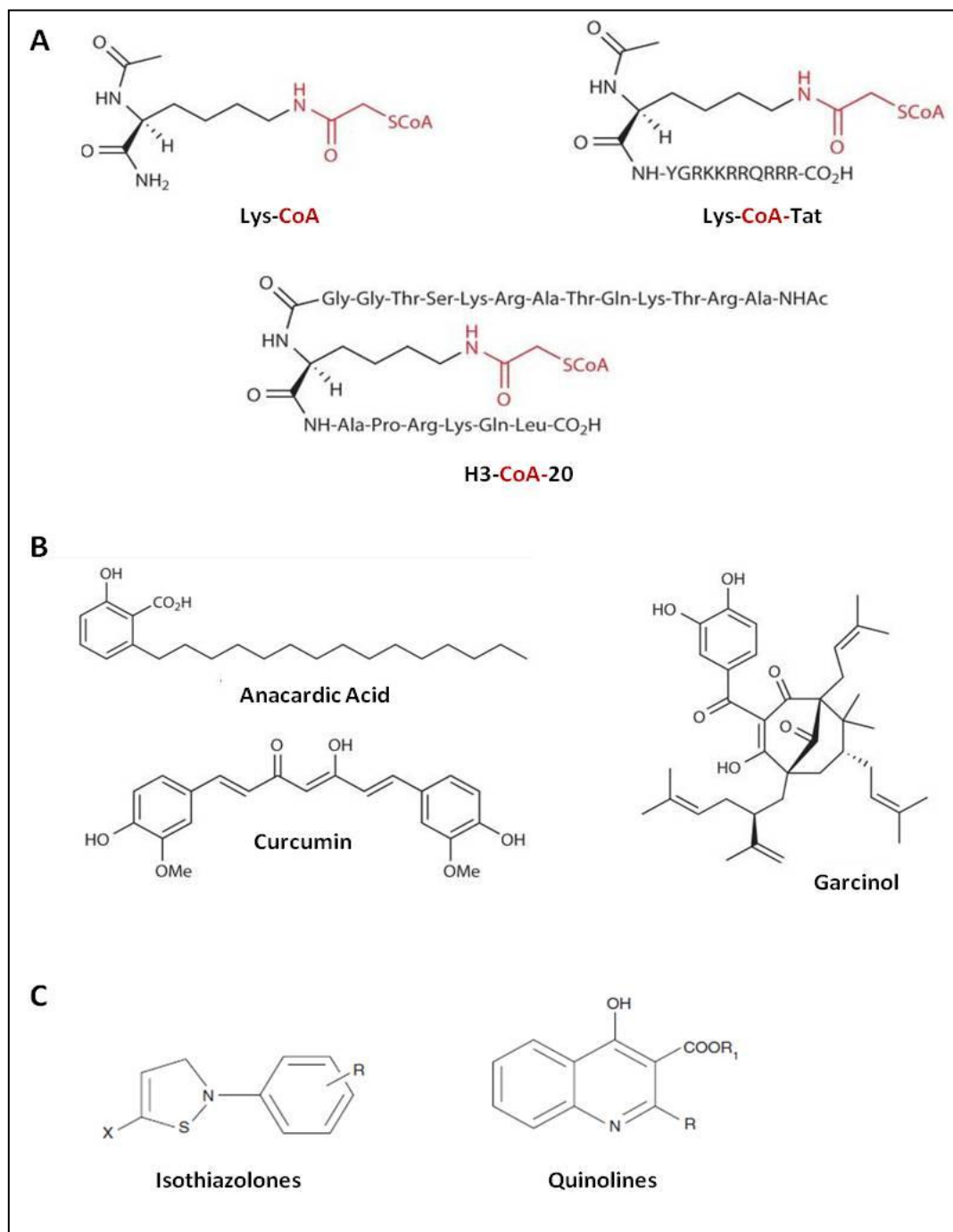
Curcumin is a yellow pigment extracted from the root of the turmeric herb *Curcuma longa* and has long been recognised for its medicinal properties. With

regards to its effects on KATs, curcumin has been shown to be a non-competitive inhibitor of p300 activity *in vitro* (Dekker and Haisma, 2009). It is cell-permeable allowing the use of the KATi in an *in vivo* context to specifically modulate KAT activity (Balasubramanyam et al., 2004). Previously, Balasubramanyam and colleagues have demonstrated that curcumin has no effect on other histone-substrate enzymes such as PCAF, HDAC1 and methyltransferases (Balasubramanyam et al., 2004). However, unpublished data shows that the HAT activities of MYST family members, MOZ and TIP60, are also inhibited by curcumin (Collins, Heery - unpublished). A promising observation of curcumin activity is that it is capable of inducing apoptosis in cancer cells without cytotoxic effects on healthy cells (Tourkina et al., 2004). In addition, curcumin has been found to possess antioxidative and anti-inflammatory properties (Selvi and Kundu, 2009) and it is due to these pharmacological properties that curcumin is under clinical study for many human pathologies (Manzo et al., 2009). In terms of searching for an exclusive KAT-specific inhibitor new products are still being actively pursued; as curcumin is a well-known antioxidant, it is therefore likely to interfere with other cellular pathways. The bioavailability of curcumin is also very low and further research needs to be carried out in order to improve this factor. Yet, the simple structure of this molecule allows the easy access to derivatives and the potential to overcome this shortcoming.

Garcinol is an isoprenylated benzophenone isolated from *Garcinia indica* fruit rind. It was shown to be a highly potent competitive inhibitor of p300 and PCAF, which inhibits histone acetylation *in vivo* and induces apoptosis by binding to the acetyl-CoA ligand site (Arif et al., 2009). Favourably, garcinol is highly permeable to cultured cells however, its poor solubility and non-specific nature render it cytotoxic.

Conversely, this compound provides a foundation for medicinal chemists to synthesise more potent, selective and less toxic analogues. Indeed, optimisation of such leads has yielded some promising results. Mantelingu *et al* have modified garcinol to create isogarcinol (IG) and 14-methoxyIG (LTK-14) derivatives (Mantelingu et al., 2007). It was found that the LTK-14 compound specifically inhibited p300 activity by a noncompetitive mode of inhibition towards both the histone and acetyl-CoA substrates and additionally no longer affected the activity of PCAF (Arif et al., 2009). In the presence of 10  $\mu$ M LTK-14, p300-mediated acetylation of histones H3 and H4 were equally inhibited up to 85-90 % compared to control. In addition, the garcinol derivative inhibited histone acetylation in HIV-infected cells and further, impaired the multiplication of HIV (Mantelingu et al., 2007). The revelation of highly specific inhibitors holds not only enormous therapeutic potential, but also provides exciting opportunities to probe into KAT function.





**Figure 1.7. Chemical structures of known KAT inhibitors. (A) Bi-substrate analogues. (B) Natural compounds. (C) Small molecule synthetic inhibitors.**

High-throughput screening of synthetic libraries has also yielded candidate chemical probes for KATs (figure 1.7C), including  $\alpha$ -Methylene- $\gamma$ -butyrolactones, isothiazolones and quinolines (Manzo et al., 2009). Isothiazolones have been found to inhibit p300 and PCAF activity and also possess antiproliferative properties against cancer cell lines. Both  $\alpha$ -Methylene- $\gamma$ -butyrolactone and quinoline inhibitors have been shown to inhibit the catalytic mechanism of GCN5 (Dekker and Haisma, 2009) (Manzo et al., 2009).

In a recent study by Buczek-Thomas a novel class of KAT inhibitors was introduced. They report that glycosaminoglycans (GAGs), with emphasis on heparin and heparan sulphate proteoglycans (HSPGs), are potent inhibitors of p300 and PCAF activities *in vitro*. The PCAF binding heparin was shown to act as a competitive-like inhibitor, causing an approximate 50-fold decrease in binding affinity of PCAF for its natural substrate histone H4 (Buczek-Thomas et al., 2008). This poses the theory that GAGs act as natural inhibitors of KAT enzymes, thus regulating histone acetylation and accordingly modulating cell function.

Research into effective KATi is still in its infancy, especially with regards to searching for inhibitors with a high degree of enzyme specificity and potency towards their targets. However, promising results are emerging as some inhibitors are in the process of clinical trials. The main focus of inhibitor production is currently aimed at the p300/CBP and GCN5/PCAF family members, where the already characterised and available inhibitors can serve as the templates for enhancement by structure-based and rational inhibitor design. Currently, inhibitors that target MYST family members are yet to be discovered; possibly due to their broad range of substrate specificities and functions. However, unpublished data from our group suggests that

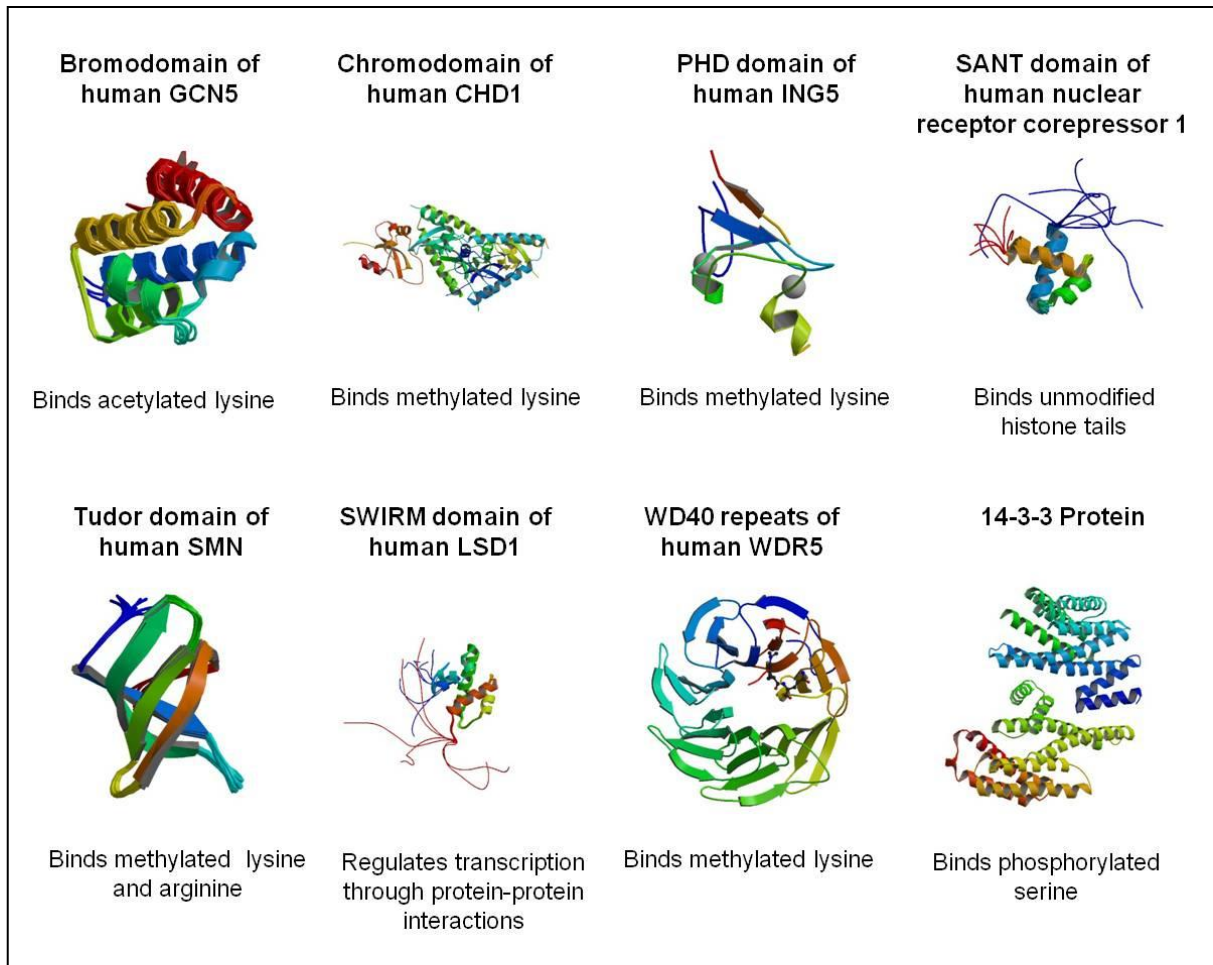
both the previously characterised curcumin and garcinol inhibitors can also affect the activities of MYST proteins (Collins, Heery – unpublished). Efforts to identify MYST-specific inhibitors are ongoing.

### 1.9. Readers of the Histone Code

Mechanistic understanding of how the plethora of histone PTMs are deciphered and translated into specific functional outputs, such as transcription, DNA repair and epigenetic silencing, has flourished over the last decade. At present, there are two non-mutually exclusive models dictating how histone PTMs may generate critical structural alterations in the chromatin polymer; the direct and effector-mediated models. The direct model implicates histone PTMs in the direct alteration of chromatin compaction, whereby PTMs such as acetylation and phosphorylation serve to disrupt favourable histone-DNA and internucleosomal contacts. Conversely, the emerging effector-mediated concept proposes that histone PTMs are ‘read’ by protein modules and their associated complexes, facilitating significant downstream events via the recruitment or stabilisation of non-histone binding effector proteins and machinery (Seet et al., 2006).

Numerous conserved protein domains have been identified (Figure 1.8) and these modules have been shown to specifically bind covalent moieties depending on both their modification state and their position within a histone sequence (Taverna et al., 2007). Emerging research underscores the importance of the multivalency concept in chromatin biology, where PTMs are read combinatorially as a ‘language’ rather than a strict ‘code’. Early studies focus on the characterisation of an individual

histone reader module binding to a single modification however more recently, research is advancing in order to define the readout of more complex epigenetic signatures. The coexistence of chromatin modifications on one or two tails of the same nucleosome, adjacent or discontinuous nucleosomes can be cooperatively or synchronously read by single binding modules or by the combinatorial engagement through effectors harbouring several putative interaction domains (Ruthenburg et al., 2007b).



**Figure 1.8. Chromatin-binding domains: readers and interpreters of the histone code.**

A number of chromatin-binding modules have been found associated with a wide array of effector proteins. Illustrated here are the various domains with their associated structures (Protein Data Bank accession numbers: bromo, 1F68; chromo, 2B2Y; PHD, 3C6W; SANT, 2EQR; Tudor, 1G5V; SWIRM, 2COM; WD40, 2G9A; 14-3-3, 2C1J), as well as which PTMs the domain recognises.

### 1.9.1. Recognition of acetyl-lysine modifications

Bromodomains (BD) are protein modules that interpret acetyl-lysine marks and are predominantly found in nuclear chromatin-associated proteins, such as KATs and subunits of remodelling complexes, in which they conduct crucial roles in transcriptional activation and chromatin remodelling. The majority of BDs utilise a similar method to recognise acetyl-lysine (Kac) residues and the BD residues engaged in Kac recognition are among the most highly conserved amino acid residues of the large BD family (Zeng et al., 2008b). Acetyl-lysine binding can be illustrated with the NMR-determined structure of Gcn5 BD as an example. The acetyl-lysine inserts into a deep and narrow binding pocket lined with aromatic and hydrophobic residues. The acetyl modification is specifically anchored through a hydrogen bond involving its acetyl carbonyl group and the amide nitrogen of a conserved asparagine residue. Further, a network of water-mediated intermolecular hydrogen bonds at the base of the cleft contributes to binding (Figure 1.9 (A))(Owen et al., 2000). Additional contacts are also made between peptide residues flanking the Kac and the amino acids residing at the entrance of the binding pocket. These interactions serve to reinforce the binding of the Kac of the target sequence. Structurally these residues compose highly flexible loop regions, which among BDs vary greatly in sequence and confer to their distinct histone-binding selectivity (Zeng et al., 2008b).

In analyses of the ligand binding sites outside of the conserved scaffold, CBP and PCAF BDs were found to adopt two distinct consensus histone recognition patterns. PCAF was found to favour acetylation sites with a hydrophobic and positively charged or aromatic residue at position Kac+2 and Kac+3 respectively;

whereas CBP preferred a consensus of bulky hydrophobic residues at Kac+1 and Kac+2, a positively charged residue at Kac-1 and an aromatic residue at Kac-2 (Zeng et al., 2008b).

The effector-mediated recognition of single PTMs is well characterised (Taverna et al., 2007), however how the cell conducts the combinatorial interpretation of histones or nucleosomes harbouring multiple modifications is still ambiguous. Recent progress, utilising structural and biophysical methods, have determined a number of distinct mechanisms involved in multivalent binding. Firstly, the mouse TAF1 homologue Brdt contains two bromodomains each with different PTM binding specificities. The single Brdt BD1 is responsible for selectively recognising hyperacetylated histone H4 harbouring two or more acetylation marks. The crystal structure of the BD1 in complex with a H4K5/8 di-acetylated peptide illustrates how two acetyl-lysine residues cooperatively bind with a single binding pocket. Binding entails the prototypical recognition of H4K5ac as described previously for the GCN5 BD, but conversely, a single amino acid substitution widens the Brdt BD1 binding pocket from the keyhole-like state of the GCN5 pocket (Moriniere et al., 2009). This wider pocket favourably accommodates both acetyl-lysines recognising the H4K8ac predominantly through hydrophobic interactions. Specificity for H4K8ac also arises structurally as the positive charge of an unmodified H4K8 residue would be adverse to the apolar environment (Moriniere et al., 2009). In addition, the BRD2-BD2 has been structurally characterised bound to the di-acetylated H4K5/12 peptide. In contrast to BD1, a single H4K5/12ac peptide interacts with two molecules of BRD2-BD2 molecules simultaneously: the K5ac residue to one molecule and the K12ac residue binds to another (Umehara et al., 2010). These results provide a structural

basis for two distinct methods of bivalent histone PTM recognition by a single bromodomain.

An alternative mechanism involves the multivalent PTM binding by linked binding modules. The tandem bromodomains of TAF1 exhibit stronger binding affinities to a di-acetylated histone H4 peptide in comparison to mono-acetylated histone H4 tails. Each BD engages one acetyl-lysine residue resulting in a cooperative binding effect between the tandem effector modules (Jacobson et al., 2000). The majority of studies thus far have been conducted with peptide surrogate substrates, yet little is known about how combinations of PTMs are recognised in a more physiologically relevant nucleosomal context. Ruthenburg *et al.* characterised the multivalent interactions of the bromodomain and PHD domain transcription factor (BPTF) PHD-Bromodomain cassette using both peptide surrogates and semi-synthetic histones reconstituted into mononucleosomes. The BPTF PHD domain is known to interact with H3K4me2/3 (Li et al., 2006) and via systemic screening with modified peptides, BPTF PHD-BD exhibited specificity for three acetyl-lysines H4K12/16/20ac (Ruthenburg et al., 2011). Biophysical and structural analysis revealed simultaneous bivalent binding by both modules, where binding was augmented for nucleosomes bearing both the H4K16ac and H3K4me3 modifications over mononucleosomes harbouring only H3K4me3. Furthermore, the bivalent nature of this interaction appears to enhance the selectivity of the BPTF PHD-BD cassette towards H4K16ac in combination with H3K4me3 binding at the mono-nucleosomal level, despite a lack of discrimination for the H4Kac peptide surrogates (Ruthenburg et al., 2011). These studies emphasise the importance in deciphering the nucleosomal

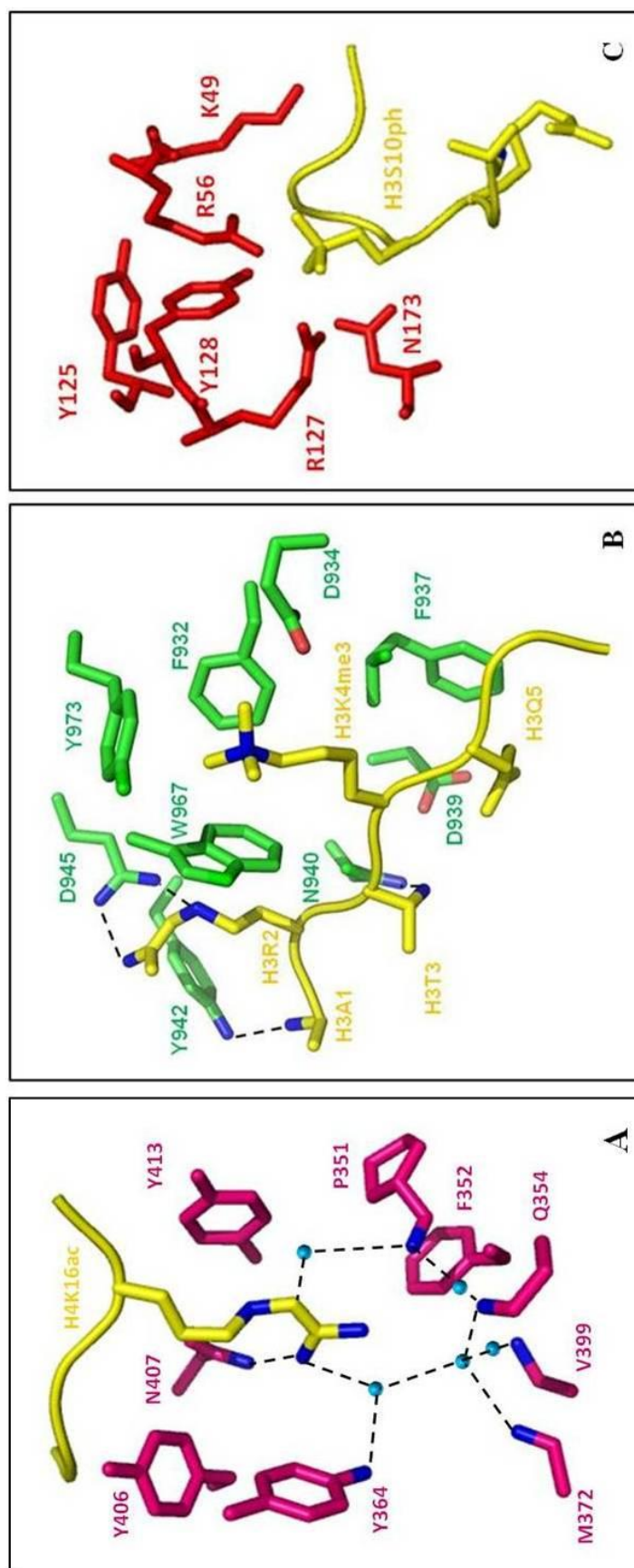


patterning of covalent marks, which dictate vital chromatin associations and downstream biological events.

### 1.9.2. Recognition of methyl-lysine/arginine modifications

Two general classes of protein domains named the PHD fingers and Royal superfamily have evolutionary converged to bind methyl-lysine/arginine residues. The greater number of modules dedicated to methylation rather than acetylation interpretation is perhaps a result of the increased combinatorial complexity of the methylation modification. Additionally, whereas lysine charge is neutralised by acetylation all methylated states of lysine are predicted to carry a positive charge at physiological pH. With the addition of each methyl group, hydrophobicity and the cationic radius of the methylammonium group increases and consequently its ability to donate hydrogen bonds decreases (Taverna et al., 2007). The different methylation states thereby diversely alter the physiochemical properties of lysine residues, requiring state-specific readout by multiple effector domains.

The recognition and interaction of methyl-lysine residues is mediated through contacts made between the methyl-ammonium moiety and aromatic residues as well as one or more acidic side chains within the binding module. Aromatic residues form a 'cage' around the PTM, creating an aromatic  $\pi$  system and a negative potential field, which promotes the engagement of the quaternary ammonium moiety with favourable electrostatic cation- $\pi$  interactions. For recognition of the lower methylation states, steric exclusion and hydrogen bonding are progressively more significant (Ruthenburg et al., 2007a).



**Figure 1.9. Chromatin-binding modules: structural interpretation of the histone code.** Schematic representation of the main interactions involved in the recognition of histone post-translational modifications. All figures were adapted from PDB structural information using the PyMOL™ 2006 software. (A) Recognition of H4K16ac marks by the Gcn5 bromodomain (PDB 1E6I). (B) Recognition of H3K4me3 modification by the double tudor domain of JMJD2A (PDB 2GF7). (C) Recognition of the H3S10ph mark by the 14-3-3δ/ζ isoforms (PDB 2C1J). Blue circles and dashed lines represent water-mediated hydrogen bonds.

The Royal superfamily members, encompassing chromo- and Tudor domains, Malignant Brain Tumour (MBT) and WD40 repeats utilise two recognition modes to engage methylated histone marks; a surface-groove mode for higher methylation states or cavity-insertion for complexes of lower methylation states. In the surface-groove recognition mode, the binding pockets have less stringent preferences for specific methylation states and thereby tend to be wider and more accessible; allowing the bulky methyl-lysine side chain to extend along a protein surface groove (Taverna et al., 2007). Examples of such a recognition mode include the binding of H3K4me3 by the double tudor domain of the histone demethylase JMJD2A (Huang et al., 2006)(Figure 1.9(B)) and the recognition of H3K9me3 by the HP1 chromodomain (CD) (Nielsen et al., 2002) (Figure 1.13).

The MBT pocket of the tumour suppressor protein L3MBTL1 binds both mono- and di-methylated lysine residues within an aromatic cage via the cavity insertion method (Min et al., 2007). In this case the methyl-lysine is buried deep within the protein cleft, thereby discriminating against the higher degrees of histone lysine methylation by size exclusion. Furthermore, in addition to binding via a cage of aromatic residues, lower methylation state interpretation begins with the formation of a direct hydrogen bond and salt bridge between methylammonium group and the carboxylate of an acidic residue lining the aromatic pocket. Thus, trimethylation of lysine residues would be undesirable by a steric repulsion between the acidic side chain and the additional methyl group (Min et al., 2007). At present, knowledge of methyl-arginine effectors is sparse due to a lack of structural data and consequently, insights into the molecular mechanisms underlying the recognition pattern. However, recent studies show that arginine methylation antagonises the binding of other

effector modules. For example, structural characterisation reveals that the H3K4me3 binding domains, including the ING2 PHD finger and WD40 repeats of WDR5, make critical contacts with the unmodified H3R2 residue. Results show that effector proteins are sensitive to H3R2 methylation, which disrupts complex interactions and antagonises binding to the N-terminal tail of histone H3 (Iberg et al., 2008). These findings underscore that it is the combinatorial readout of PTMs that govern the final outcomes dictated by the reader modules.

### 1.9.3. Recognition of phospho-serine/threonine modifications

Multiple serine, threonine and tyrosine histone residues can be modified by phosphorylation, however only a small number of phospho-histone effectors have been identified, despite the description of numerous phospho-binding modules for non-histone proteins. The mammalian 14-3-3 proteins represent an abundant and well-conserved family of phospho-specific binding proteins and structure-function studies have classified 14-3-3 isoforms that bind N-terminal H3 tails in a phosphorylation-dependent manner (Macdonald et al., 2005). Phosphorylation affixes a bulky and negatively charged phosphate group onto the hydroxyl group of a side chain, thus increasing the ion-pairing and hydrogen-bonding abilities of these residues (Taverna et al., 2007). The crystal structure of the 14-3-3 $\zeta$  isoform showed that the phosphorylated (S10) peptide binds in an extended conformation spanning the entire 14-3-3 binding cleft. This extended conformation is stabilised by intermolecular hydrogen bonds between the side chains of 14-3-3 asparagine residues and the backbone carbonyl and amide groups of the H3 peptide (Macdonald et al.,

2005). The S10ph group is bound in a pocket lined with highly basic amino acid residues, which act in neutralising the negatively charged phosphate moiety. Further, the expulsion of solvent and the presence of aromatic residues contribute to the charged character of the binding cleft (Figure 1.9(C)) (Macdonald et al., 2005). Regarding the crosstalk between and influence of PTMs on the same histone tail, it was found that acetylation of K9 or K14 had a minimal effect on binding of the 14-3-3 $\zeta$  motif with the phospho-serine residue. A H3K9ac/S10ph peptide binds in a partially extended conformation with a kink, causing the side chain of K9ac to fold back and form a hydrogen bond to the backbone amide of A7, thus not disturbing the 14-3-3-S10ph interaction (Macdonald et al., 2005).

### 1.9.4. Recognition of unmodified amino acid residues in histones

An important concept is that complexes also interact with unmodified histone tails through SANT, WD40 repeats and PHD domains. This underscores the importance of the residue context, where the same residue in the absence or presence of a PTM, can be used to promote different chromatin associations or, in other instances weaken or block these interactions (Taverna et al., 2007). This thereby establishes the significance in the ability to also recognise residues in the absence of PTMs. Examples of the molecular recognition of unmodified residues by PHD fingers will be outlined in the subsequent section. In general, all effector binding modules form associations with unmodified histone residues. In addition to the primary recognition of the target PTM, an extensive network of interactions between the histone peptide backbone and binding pocket residues is required for the

stabilisation and efficient recognition of the histone PTM. This concept is perfectly illustrated by the previously described WDR5, a seven-WD40-repeat, H3K4me3 binding protein that forms critical interactions with the peptide backbone residue H3R2. Indeed, crystal structures show H3R2, located on the exposed tail of histone H3, contacts the centre of the WDR5  $\beta$ -propeller structure. Once H3R2 is methylated by the PRMT6 HMT, binding and docking to its target H3K4me3 PTM is prohibited (Iberg et al., 2008).

### 1.10. The PHD Finger: A Versatile Protein-Interaction Domain

Almost two decades ago, Schindler *et al.* first identified the PHD finger motif within an Arabidopsis homeodomain protein. The presence of eight regularly spaced cysteine/histidine residues was observed between two plant homeodomain proteins and such sequence similarity is concomitant with other metal-binding domains such as the RING (Really Interesting New Gene) finger. This novel motif was denoted PHD-finger (Plant HomeoDomain) (Schindler et al., 1993). Since then the PHD finger has been classified as a *bona fide* domain that is present in an array of eukaryotic proteins. PHD motifs tend to comprise between 50-80 amino acids and show a C4HC3 signature that structurally adopts a 'cross-brace' topology of the two  $\text{Zn}^{2+}$ -coordinating residues (Bienz, 2006). PHD fingers are predominantly found in nuclear proteins that function in the regulation of chromatin, suggesting that they bind to a common nuclear ligand. However, due to low sequence similarity and structural flexibility among PHD finger domains, it is emerging that multiple biological ligands can be bound.

### 1.10.1. PHD fingers as chromatin binding modules

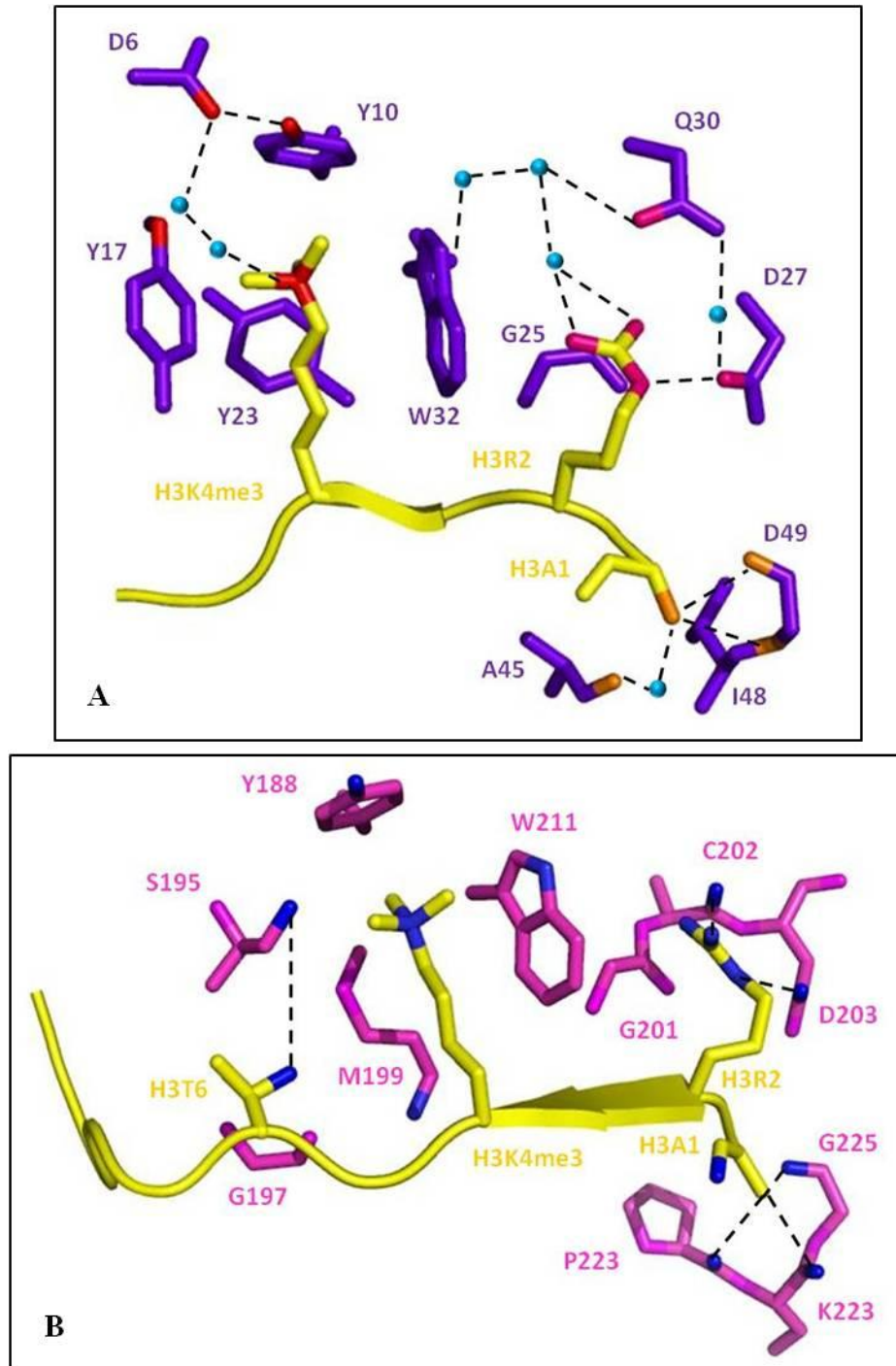
The original breakthrough of a function for PHD domains in gene transcription was assigned to the recognition of lysine-methylated histone H3 (Li et al., 2006). The PHD finger is now a well-established chromatin-binding module that recognises and interprets both methylated, acetylated or the unmodified states of histone lysine residues. Further, recent advancements in the structural and functional characterisation of PHD fingers underscore their functional versatility as epigenome readers and reveal a refined histone sequence reading capability that is governed by the cross-talk between different histone PTMs (Sanchez and Zhou, 2011).

The first solution structure of a PHD finger bound to a histone ligand was that of the human BPTF, the largest subunit of the ATP-dependent chromatin-remodelling complex, nucleosome remodelling factor (NURF) (Li et al., 2006). The BPTF C-terminal PHD finger was shown to specifically bind the histone H3 tail when H3K4 is tri-methylated and H3R2 is unmodified, discriminating against the unmodified and mono-methylated H3K4 counterparts. The BPTF-PHD structure also set a precedent for the main characteristics required for the readout of H3K4me3 by PHD fingers. The binding of H3K4me3 occurs via an aromatic cage (Figure 1.10(A)), where the trimethyl ammonium moiety is stabilised by electrostatic cation- $\pi$  interactions, like those observed in members of the Royal superfamily (section 1.9.2). Importantly, R2 and H3K4me3 fit into two adjacent surface pockets divided by an invariant tryptophan group of the PHD finger (Figure 1.10(A)). The observed positioning of non-adjacent R2 and K4me3 side chains provides a molecular basis for H3K4me3 site specificity (Li et al., 2006). Subsequently determined structures of other PHD

fingers bound to H3K4me3 peptides revealed that the aromatic cage varies, consisting of a combination of two to four aromatic and hydrophobic residues (Sanchez and Zhou, 2011). For instance the PHD finger of ING5, an integral subunit of active MOZ/MORF and HBO1 complexes, essentially utilises the same basis of recognition specificity as the BPTF PHD-H3K4me3 complex. However, the H3K4me3 binding pocket comprises only a partial aromatic cage, whereby the hydrophobic character of the two aromatic residues is supplemented by the side chains of serine and methionine residues (Figure 1.10(B)) (Champagne et al., 2008).

A separate group of PHD fingers, including those of BRAF35-HDAC (BHC80) (Lan et al., 2007), autoimmune regulator (AIRE) (Org et al., 2008) and tripartite motif-containing protein 24 (TRIM24) (Tsai et al., 2010), have evolved to bind unmodified H3K4me0. The crystal structure of the BHC80 PHD finger, a component of the LSD1 corepressor complex, bound to an un-methylated peptide revealed the structural basis of the recognition of H3K4me0. Similar to its methylated counterparts, the N-terminus of the H3 peptide is recognised by a hydrogen bond cage formed by three backbone carbonyls (Figure 1.11). H3K4me0-binding PHD fingers lack an aromatic cage and binding to the free epsilon amino group is achieved by an arrangement of acidic and hydrophobic residues instead. The H3 peptide binding site is defined by a methionine, which is interposed between H3R2 and H3K4 and a side chain of an aspartate residue, which inserts between the side chains of H3K4 and H3R8, forming an electrostatic bridge between the two (Figure 1.11) (Lan et al., 2007).



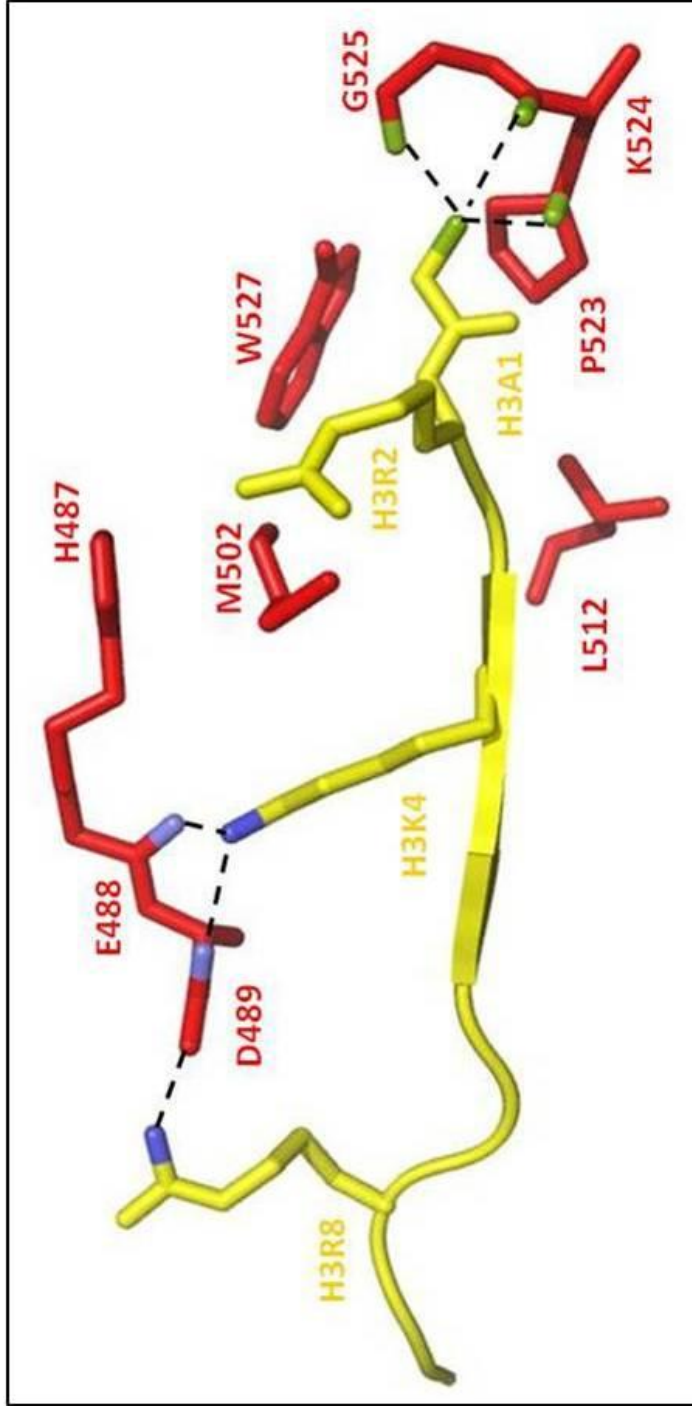


**Figure 1.10. Recognition of the H3K4me3 modification by the PlantHomeo Domain.** Schematic representation of the main interactions involved in the recognition of histone H3K4me3 by (A) an aromatic cage of BPTF PHD finger (PDB 2F6J) or (B) a partial aromatic cage of the ING5 PHD finger (PDB 3C6W). Blue circles and dashed lines indicate water-mediated hydrogen bonds. This figure was adapted from PDB structural information using the PyMOL™ 2006 software.

In addition, H3K4 forms a hydrogen bond with a main chain carbonyl oxygen of a glutamic acid residue (Figure 1.11) (Lan et al., 2007). The presence of these acidic side chains further confer binding specificity by creating an unfavourable environment for bulky cationic methyl groups, thus resulting in the steric exclusion of the equivalent H3K4me<sub>2/3</sub> modifications (Lan et al., 2007). Recently, the PHD finger of AIRE has been shown to be sensitive to H3R2 methylation and therefore reads a completely unmodified H3 tail (H3R2me<sub>0</sub> + H3K4me<sub>0</sub>); a property unique among the H3K4me<sub>0</sub> histone readers. A salt bridge is formed between an aspartate residue and the unmodified H3R2, which is disrupted upon methylation of H3R2 (Chignola et al., 2009). This represents the readout of a histone PTM signature by a PHD finger and details the interpretation of negative cross-talk that exists between the unmethylated H3K4 and methylated H3R2 modifications.

Details from the few examples provided show that PHD fingers are capable of exerting such multifaceted functional flexibility as chromatin readers. However, reading of histone modification status is enhanced by the cooperative capacity achieved through the presence of two or more PHD fingers (or other reader modules) in the same protein or within a protein complex. The simultaneous bivalent interaction of the BPTF PHD-BD cassette to H3K4me<sub>3</sub> and H4K16ac described above (section 1.9.1) (Ruthenburg et al., 2011) illustrates clearly an example of intramolecular cooperative histone PTM binding. It is also emerging that multisubunit complexes utilise multiple binding modules to guide the appropriate targeting and association of the complex to chromatin. The HBO1 HAT complex contains three PHD fingers in two different subunits; tumour suppressor proteins ING4/5 and JADE 1/2/3. Each of the PHD fingers was found to interact with the N-

terminal tail of histone H3, but with different specificity towards its methylation status (Saksouk et al., 2009). In this example the combinatorial action of all three PHD fingers is required for the regulation of chromatin association and substrate specificity of HBO1 complexes (Saksouk et al., 2009).



**Figure 1.11. Recognition of the H3K4me0 modification by the BHC80 PHD finger.** Schematic representation of the main interactions involved in the recognition of unmodified histone H3. The N-terminal H3A1 is accommodated by a hydrogen bond cage. H3K4me0 binding is achieved by acidic and hydrophobic residues, whereby BHC80 E488 forms a hydrogen bond with the H3K4 peptide residue and the BHC80 D489 forms an electrostatic bridge between H3K4 and H3R8 histone peptide residue. Dashed lines indicate hydrogen bonds. This figure was adapted from PDB structural information (2PUY) using the PyMOL™ 2006 software.

### 1.10.2. Tandem PHD fingers as histone acetyl-lysine sensors

Many proteins, especially those that function in the regulation of chromatin, have been shown to contain single or multiple PHD fingers. However, within the genome only five proteins harbour a double PHD finger domain where the PHD fingers are arranged in tandem; namely MOZ and MORF MYST family histone acetyltransferases and members of the d4-protein family, DPF1/2/3.

DPF1 (Neuro-d4), DPF2 (ubi-d4/requiem) and DPF3 (Cerd4) are evolutionary conserved members of the D4, zinc and double PHD finger family characterised by an N-terminal 2/3 domain, a C2H2-type zinc finger and C-terminal tandem PHD fingers. Mouse Dpf1 expression is predominantly restricted to the brain and neuronal systems, where it is thought to function as a neurospecific transcription factor in neuronal development through the regulation of cell survival (Lessard et al., 2007). DPF2 is ubiquitously expressed and functions as a transcription factor necessary for the apoptotic response following the deprivation of survival factors (Gabig et al., 1994). At a molecular level, DPF2 has recently been shown to act as a nuclear receptor-selective corepressor for  $ERR\alpha$  by associating with both acetylated H3 and HDAC1 in myeloblast cells (Matsuyama et al., 2010). DPF3 is a muscle expressed member of the d4-family. The human *DPF3* gene gives rise to two splice variants; DPF3a and DPF3b. DPF3a encodes a truncated protein containing a single C-terminal PHD finger, while DPF3b harbours a C-terminal DPF (Lange et al., 2008). *In vivo* knockdown of *dpf3* reveals the protein as a key regulator of heart and skeletal muscle development (Lange et al., 2008).

All three DPF proteins were found to act as novel components of the ATP-dependent BAF chromatin remodelling complex, where their tandem PHD fingers function as novel readers of histone H3/ H4 acetylation and methylation (Lange et al., 2008). Thus, remarkably the DPF family exhibit the first PHD domain that recognises acetylated lysine residues; a property originally assigned to the bromodomain only.

The tandem PHD finger of DPF3b was shown to bind acetylated H3K14 (Lange et al., 2008) and later, the three-dimensional structure of acetylated histone binding by DPF3b was solved using NMR methods (Zeng et al., 2010). The integrated tandem PHD finger functions as one functional unit in the sequence specific recognition of the acetyl-lysine residue, whereby PHD1 is responsible for the binding of the H3K14ac residue and PHD2 binds the H3R2-H3K4 region (Zeng et al., 2010). The H3K14ac hydrophobic pocket in DPF3b-PHD1 is completely novel. Acetylated H3K14 intercalates into a hydrophobic binding pocket and clustered at the rim of this pocket are aspartate, zinc-coordinating cysteine and arginine residues. The acetyl amide of H3K14ac is positioned within hydrogen bond distance to the side chains of the aspartate and cysteine residues, whereas the acyl chain of H3K14ac interacts with the arginine side chain (Zeng et al., 2010). DPF3b-PHD2 is similar to the H3K4me0 binding PHD fingers in that it shares the arrangement of acidic and hydrophobic residues in the binding pocket. Together they define the histone PTM recognition specificity, whereby the interaction of histone H3 is stimulated by acetylation of H3K14, but inhibited by tri-methylation of H3K4 (Zeng et al., 2010).

### 1.10.3. PHD fingers as protein-protein interaction domains

The PHD finger motif has also been shown to act as an independent protein interaction domain. Enhancer of Zeste (E(Z)) and Polycomblike (PCL) are both *Drosophila* Polycomb group (PcG) proteins required for the epigenetic suppression of multiple essential developmental regulatory genes including homeotic genes. Yeast two-hybrid and co-immunoprecipitation experiments revealed a direct interaction between E(Z) and PCL both *in vitro* and *in vivo*. Subsequently, mutagenesis of the two PCL PHD fingers demonstrated that this protein-protein interaction was mediated by these conserved motifs (O'Connell et al., 2001). This interaction was observed to be evolutionarily conserved as the human homologues PHF1 (PCL) and EZH2 (E(Z)) were seen to associate. The interaction was also highly specific as no binding was observed between E(Z) and the PHD domains of *Drosophila* trithorax group (trxG) proteins; proteins which also function in the transcriptional regulation of homeotic genes (O'Connell et al., 2001).

As another example, the interaction between the BCL9 and Pygopus components of the Wnt signalling pathway is mediated via the Pygopus PHD finger. Mutation of a structural zinc-coordinating cysteine and also four conserved residues at the putative PHD domain surface abolishes the interaction with BCL9 (Townesley et al., 2004). Three of these surface residues are identical amid all known Pygopus proteins, but different in comparison to other PHD fingers, suggesting that these residues define the specificity of the BCL9-Pygopus interaction (Bienz, 2006).

### 1.10.4. PHD E3 ligase activity

The closest structural relative to the PHD finger is the RING domain. The RING domain typically shows a C4HC3 signature and shares similarity with PHD domains in their structural cores and ‘cross-brace’ topology of the  $\text{Zn}^{2+}$  coordinating residues (Bienz, 2006). RING fingers are motifs commonly found in E3 ligases functioning in the ubiquitin pathway, where they bind to E2 ligases to mediate protein ubiquitination (Joazeiro and Weissman, 2000). Hence, the similarity between the two domains prompted reasoning that PHD domains may also contain this intrinsic E3 ligase activity. To date, few examples exist of PHD domains having been assigned with this function due to controversy amid the classification of specific types of zinc fingers. This controversy highlighted the need for better criteria and tools to delineate and distinguish between these motifs (Bienz, 2006). Despite this, one canonical PHD finger has now been implicated as a functional E3 ligase for SUMO molecules.

Recent studies have shown that PHD and bromodomain modules located in close proximity in chromatin-associated proteins, functionally cooperate in the combinatorial recognition of histone modifications (Ruthenburg et al., 2011). In addition, work from Ivanov *et al.* demonstrates an alternative mechanism of synergy based on the enzymatic activity of a PHD finger, directing the modification of the bromodomain. The KAP1 (KRAB-associated protein 1) corepressor mediates gene silencing through the recruitment of nucleosome remodelling and deacetylation (NuRD) complex and the histone H3K9 methyltransferase SETDB1. Subsequent histone deacetylation and H3K9 methylation stabilises the interaction of HP1 with the target genes, formulating a silent gene expression state. SUMOylation was found to



be essential for KAP1-mediated repression (Li et al., 2007b) and the two SUMO acceptor sites, crucial for such repression, were mapped to within the bromodomain (Ivanov et al., 2007). Consequent SUMOylation of the KAP1 bromodomain regulates the recruitment of chromatin modifiers required for establishing silent-gene expression states at specific chromatin loci (Ivanov et al., 2007). Ivanov *et al.* demonstrate that the PHD domain is required for both E2 enzyme Ubc9 binding and directing SUMOylation to the specific lysine residues in the bromodomain. Thus, the PHD finger functions as a unique E3 ligase directing intramolecular SUMOylation of the adjacent bromodomain. Furthermore, the NMR structure of the PHD-BD has been solved revealing a completely unique unified tandem-domain architecture that is distinct to other PHD-BD structures (Zeng et al., 2008a). The KAP1 PHD finger and BD do not exhibit typical methyl- and acetyl-lysine binding activities, but instead function interdependently to bind Ubc9 at the tandem-domain interface and catalyse SUMOylation within this tandem module to stimulate co-repression activity (Zeng et al., 2008a).

### 1.10.5. PHD fingers as nuclear phosphoinositide receptors

Phosphoinositides (PtdInsPs) are lipid signaling molecules which play critical roles in cytoplasmic signal transduction pathways. Biochemical studies have also detected PtdInsPs within the nucleus and additionally, evidence has emerged that these molecules can regulate several nuclear processes (Chen et al., 2002). However, their functions within the nucleus are ambiguous as very few binding modules for PtdInsPs have been identified. Gozani *et al.* have utilised screening techniques to identify the ING2 tumour suppressor protein as a candidate nuclear PtdInsP binding

partner. Specifically, the ING2 PHD finger motif was found to interact *in vitro* and *in vivo* with PtdIns(5)P and as a result, was observed to functionally modulate the subcellular localisation and aptitude of ING2 to stimulate p53 and p53-dependent apoptotic pathways (Gozani et al., 2003).

### 1.11. PHD Fingers as Epigenetic Effectors and Potential Drug Targets

Cancer initiation and development can arise from the inappropriate expression or silencing of gene programs, where the balance of underlying epigenetic marks are altered. Disease-related mutation or dysregulation has been identified in abundant modification ‘writer’ and ‘eraser’ enzymes, however recently an array of diseases have been associated with the dysfunction of proteins that contain chromatin recognising ‘reader/effector’ modules, markedly PHD fingers in numerous cases (Baker et al., 2008). This signifies a new category of diseases that stem from the misinterpretation of the epigenetic ‘code’ and examples of such diseases are consequently described.

Recent results have provided the first evidence showing that disruption of PTM recognition by PHD fingers can cause human disease. The Recombination Activating Gene 2 (RAG2) V(D)J recombinase mediates antigen receptor assembly, a process at the hub of the adaptive immune response. Mutation of RAG2 and consequent disruption of V(D)J recombination was found to cause Severe Combined Immunodeficiency (T-B-SCID) or the less severe ‘Omenn Syndrome’ by disrupting or impairing V(D)J recombination. T-B-SCID patients are vulnerable to infection due to a lack of functional B and T cells (Schwarz et al., 1996). Omenn Syndrome

patients suffer from autoimmune problems and chronic infections, a result of a lack of B cells and elevated levels of T cells (Marrella et al., 2008). RAG2 contains a C-terminal PHD finger which has been shown to be critical for efficient V(D)J recombination *in vivo*. More specifically, evidence demonstrates that the PHD finger recognises and binds to the H3K4me3 modification enriched at V(D)J segments ready to undergo recombination (Matthews et al., 2007). Mutational analysis revealed that a quarter of known disease-causing mutations are located within the RAG2 PHD finger and the severity with which these mutations disrupt the PHD-H3K4me3 interaction frequently correlates with the severity of the immunodeficiency disease (Schwarz et al., 1996). These results underscore the importance of the ‘reader’ PHD module in targeting and stabilising the interaction of RAG recombinase enzymatic activity to the correct gene segments awaiting V(D)J recombination.

In addition to immune diseases, mutations in the genes encoding PHD finger-containing factors are strongly linked with neurological disorders. Haploinsufficiency of CBP leads to Rubinstein-Taybi syndrome (RTS), a congenital developmental disorder characterised by retarded growth, mental functions and facial abnormalities (Petrij et al., 1995). Mutational analysis of RTS-patients revealed two novel mutations within the PHD finger of CBP, which functionally resulted in the complete loss of acetyltransferase activity towards histones and CBP itself and demonstrated diminished coactivator function for the TF CREB (Kalkhoven et al., 2003). These findings emphasise the significance of the PHD finger for *in vivo* CBP function where a reduction in CBP HAT activity, as a result of PHD finger disruption, is sufficient for the development of RTS. ATR-X (Alpha-Thalassaemia mental Retardation, X-linked) syndrome is another example of a neurological disorder and is

characterised by severe mental and growth retardation, alpha thalassaemia and seizures. The ATRX protein harbours a PHD finger at its N-terminus and to date, 26 disease-causing mutations have been identified within this domain (Argentaro et al., 2007). The molecular effect of such mutations and their causative role in disease is currently still ambiguous and under investigation, however research focusing on the function of the ATRX PHD finger will provide valuable insights into the molecular pathogenesis of ATR-X syndrome.

The mutation of PHD reader modules has also been intimately associated with the pathogenesis of solid and blood cancers, where recent results support a causal role for the deregulated PHD finger in oncogenesis. The mammalian ING family of putative tumour suppressors function at the molecular level as chromatin regulatory molecules, operating within distinct KAT and KDAC enzyme complexes. Reduced expression, somatic mutations and allelic loss of ING proteins are detected in breast cancer, gastric cancer, melanoma, glioma and head and neck squamous cell carcinoma (HNSCC) (Shi and Gozani, 2005). All ING proteins harbour an essential C-terminal PHD finger that specifically interacts with the H3K4me3 modification. Mutational screening revealed that specific mutations that abrogate H3K4me3-binding are found in tumours (Shi and Gozani, 2005), implicating that ING PHD functional readout of H3K4 is directly involved during oncogenesis.

More recent studies have reported a fundamental role for chromatin-binding PHD fingers in haematopoietic malignancies. Chromosomal translocations of the MLL gene and resultant fusion proteins, such as MLL-ENL, are frequently associated with leukaemia. In these translocations, MLL loses the highly conserved three PHD finger cassette, SET and activation domains and fuses with a partner protein (ENL)

with a transcription activation domain. The expression of the MLL-ENL fusion protein in mice leads to HSC immortalisation and the development of myeloid leukaemia; effects caused by the aberrant activation of MLL-dependent genes (Chen et al., 2008). The third MLL PHD finger binds the nuclear Cyp33 ligand which sequentially recruits HDACs to the MLL repression domain and downregulates *HOX* gene expression. With the loss of PHD3 in MLL-ENL, this repression unit is eradicated and replaced with the ENL activation domain switching the fusion protein from a repressor to a constitutive activator of MLL-dependent *HOX* genes. Consequent overexpression of MLL target genes blocks HSC commitment and orchestrates MLL-related leukaemogenesis (Chen et al., 2008).

In AML patients, chromosomal translocations fusing the transactivation activities of Nucleoporin 98 (NUP98) to a C-terminal PHD finger of JARID1A or PHF23 are clinically reported (Baker et al., 2008). These potent oncoproteins arrested haematopoietic differentiation where cells proliferate indefinitely as undifferentiated progenitors inducing AML. In this process the gain of PHD function was essential for malignancy and the minimal requirement for oncogenesis. The fusion NUP98-PHD finger specifically recognises and binds to the H3K4me2/3 modification and mutations that abolish this interaction also eradicate leukaemic transformation (Wang et al., 2009). NUP98-PHD fusions dominate over silencing factors and enforce and ‘lock’ developmentally critical loci into an active transcription state, characterised epigenetically by H3K4me3 and induced histone acetylation. Upregulation of such developmental transcription factors defines leukaemic stem cells (Wang et al., 2009). This study represents a breakthrough in understanding the outcomes of

misinterpreting the histone code, where a direct cause-effect relationship is seen between a deregulated chromatin-binding PHD finger and oncogenesis.

These studies underscore the importance of elucidating the physiological role of epigenetic ‘reader’ modules. Functional information conveys valuable insights into normal and pathological development and may accentuate chromatin ‘readers’ as potential therapeutic drug targets.

### 1.12. Lysine Methyltransferases: Structure and Function of SUV39H1

Histone lysine methylation, with the exception of H3K79, has been shown to be catalysed exclusively by the conserved SET domain family of histone methyltransferases (Qian and Zhou, 2006). The evolutionary conserved SET domain was originally identified as a common motif in the *Drosophila* position effect variegation (PEV) modifier SU(VAR)3-9 (Suppressor of variegation) (Tschiersch et al., 1994), enhancer of zeste (E(z)) (Jones and Gelbart, 1993) and trithorax (Stassen et al., 1995). Since then a large number of SET domain-containing proteins, which contain more than 700 members, have been characterised (Qian and Zhou, 2006).

The *Drosophila* *Su(var)* group of genes suppress PEV and their gene products are implicated in the formation of higher order repressive chromatin states. *Su(var)3-9* and the human homologue *SUV39H1* encode heterochromatic proteins that harbour two of the most prominent evolutionarily conserved domains found in chromatin modulators; the chromodomain, implicated in directing the protein to distinct chromatin localisations, and the SET lysine methyltransferase domain (Figure 1.12).

SUV39H1 is a SET-domain-dependent, H3-specific methyltransferase which selectively methylates lysine 9 of the H3 N-terminus; a modification intimately correlated to the establishment of heterochromatic regions (Melcher et al., 2000).

A major function of the *Su(var)* genes determines the coregulation of higher-order chromatin at centromeres and telomeres. More specifically, SUV39H1 was found to accumulate at centromeric positions during mitosis and structure-function analysis revealed that deregulated SUV39H1 interferes at multiple levels with mammalian higher-order chromatin organisation (Melcher et al., 2000). A Heterochromatin Protein 1  $\beta$  (HP1 $\beta$ ) interaction surface was mapped to the N-terminus of SUV39H1 and together with the adjacent chromodomain, directs the specific accumulation of the complex at heterochromatin and the successive modulation of a repressive chromatin structure. Further, deregulated SUV39H1 is itself redistributed and modulates the consequent localisation of endogenous HP1 $\beta$  from heterochromatic foci to a uniform, broad-specificity distribution throughout the entire nucleus (Melcher et al., 2000).

Other nuclear processes were found to be sensitive to deregulated expression levels of SUV39H1, whereby cells over-expressing SUV39H1 displayed severe defects in mitotic progression and chromosome segregation. SUV39H1-overexpressing cells displayed abnormal nuclear morphologies and severe growth retardation, characterised by a delayed progression through the G2/M cell cycle phase (Melcher et al., 2000). These cells additionally display compromised chromosome segregation potential, exhibiting an increase in the number of cells with polynuclei and mitoses with an elevated level of chromosomal bridges and lagging chromosomes (Melcher et al., 2000). Altogether, structure-function analysis has

revealed a dominant function for SUV39H1 in heterochromatin organisation, chromosome segregation and mitotic progression.

Alternative observations implicate a function for SUV39H1 in tumour suppressor pathways. A mouse knock-out experiment showed that SUV39H1 governs H3K9 methylation at pericentric heterochromatin regions and induces a specialised histone methylation pattern that differs from the expansive H3K9 methylation present at other chromosomal regions (Peters et al., 2001). Furthermore, *Suv39h*-deficient mice exhibit chromosomal instabilities that are correlated with an elevated tumour risk and aberrant chromosome interactions during meiosis; defining a crucial role for SUV39H1-mediated pericentric H3K9 methylation in protecting genome stability (Peters et al., 2001). More recently, a SUV39H1-dependent tumour suppressor mechanism has been shown to be induced by oncogenic Ras. In response to oncogenic Ras, SUV39H1 induces H3K9me-mediated heterochromatin formation and senescence of retinoblastoma (Rb) growth-promoting genes (Braig et al., 2005). SUV39H1 deficiency blocks *ras*-induced senescence and permits the formation of aggressive T cell lymphomas (Braig et al., 2005), emphasising the HMTs tumour suppressive potential.

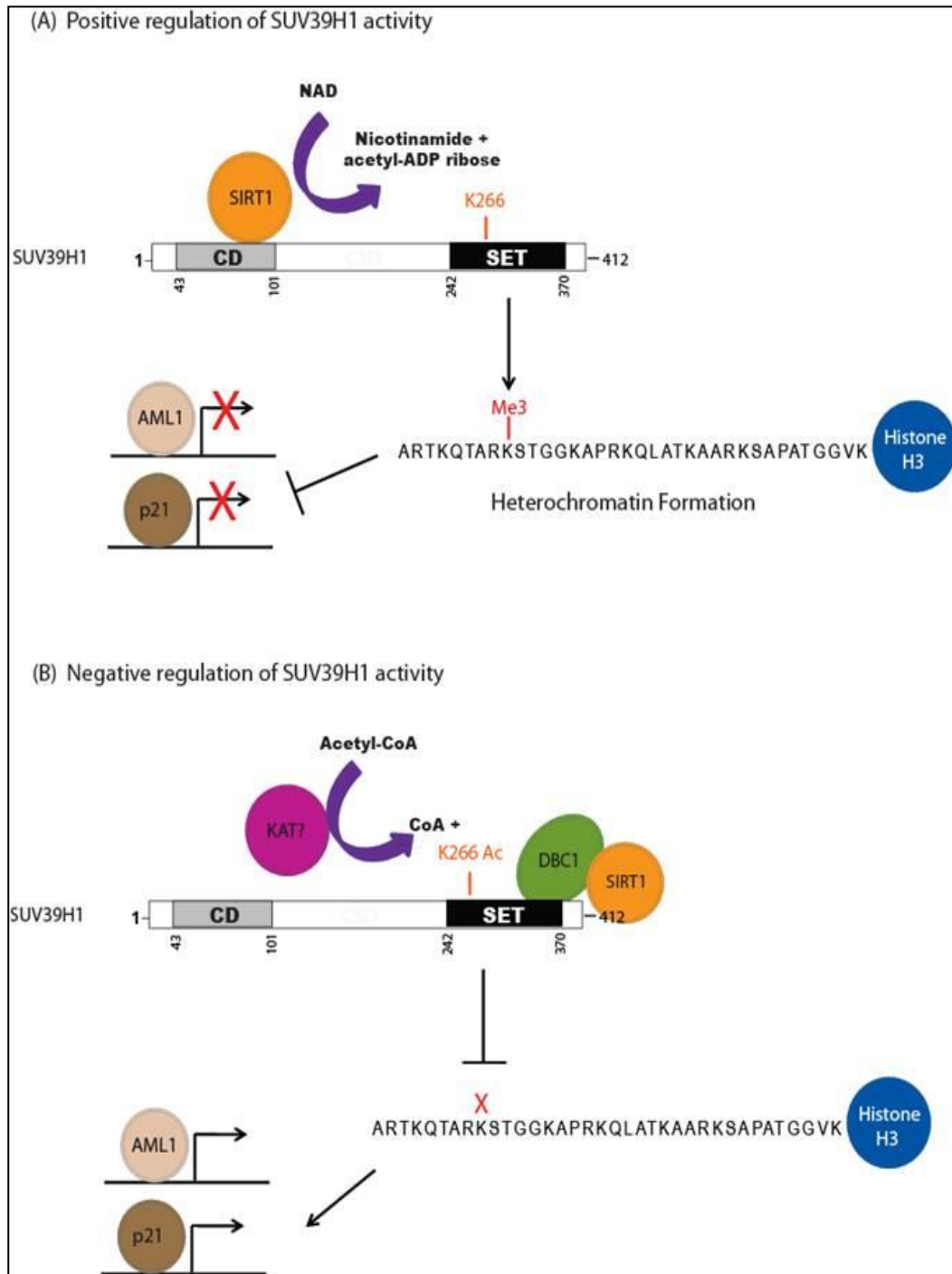
SUV39H1 has also been implicated in the epigenetic control of gene expression, where it acts as a corepressor at a number of gene promoters including AML1 and p21<sup>WAF1</sup>. AML1 forms a complex with SUV39H1 (Chakraborty et al., 2003) and histone deacetylases (Reed-Inderbitzin et al., 2006) to repress transcription at AML1-dependent target genes. By binding to AML1, SUV39H1 abrogates promoter transactivation by AML1. SUV39H1 exerts its repressor functions by interacting with the DNA binding domain of AML1 and thereby, reducing the DNA-



binding affinity of AML1. Further, coexpression of the two proteins dissociates the defined nuclear structure of AML1 from a net-like structure into a diffuse nuclear stain (Chakraborty et al., 2003). In addition, SUV39H1 acts in the negative regulation of p21, a major cell cycle regulator in the response to DNA damage, cellular senescence and tumour suppression. SUV39H1 and COUP-TF-interacting protein 2 (CTIP2) are recruited to the p21 gene promoter, SUV39H1 induces the methylation of H3K9 at the promoter region and both functionally cooperate to silence p21 gene transcription (Cherrier et al., 2009).

Sirtuin 1 (SIRT1) is an  $\text{NAD}^+$ -dependent deacetylase that specifically deacetylates H4K16 which as a consequence, through an unknown mechanism leads to elevated levels of H3K9me3 (Vaquero et al., 2004). Vaquero *et al.*, go on to demonstrate a functional link between the chromatin modifiers SIRT1 and SUV39H1, where SIRT1 positively regulates the activity of SUV39H1 during heterochromatin formation (Vaquero et al., 2007). SUV39H1 directly interacts with SIRT1 through its N-terminal chromodomain and is itself a target for SIRT1 deacetylase activity. SUV39H1 activity is regulated by acetylation at residue K266 positioned within its catalytic SET domain and deacetylation of this residue by SIRT1 augments SUV39H1 activity, resulting in elevated levels of heterochromatin-related H3K9me3 (Figure 1.12(A)) (Vaquero et al., 2007). More recently, SUV39H1 activity has been shown to be negatively regulated through the interaction with the product of the deleted in breast cancer gene (DBC1) (Li et al., 2009). DBC1 disrupts the SUV39H1-SIRT1 regulatory complex by specifically interacting with both proteins. DBC1 uses the same N-terminal domain to bind SIRT1 and SUV39H1 concomitant with its ability to disrupt the complex interaction rather than forming a trimeric

complex (Li et al., 2009). Furthermore, DBC1 binds to the catalytic SET domain of SUV39H1 abrogating its activity and ability to methylate histone H3 (Figure 1.12(B)) (Li et al., 2009).



**Figure 1.12. Schematic representation of the (A) positive and (B) negative regulation of SUV39H1 activity.** (A) SIRT1 interacts with and deacetylates lysine 266 of SUV39H1, which in turn causes elevated levels of SUV39H1 activity and H3K9me3 modification. The active form of SUV39H1 acts as a transcription repressor at gene promoters including AML1 and p21. (B) SUV39H1 activity is impaired by acetylation at lysine 266 and the binding of DBC1 to the SET domain, which both inhibit its ability to methylate histone H3.

### 1.13. Structure and Function of HP1

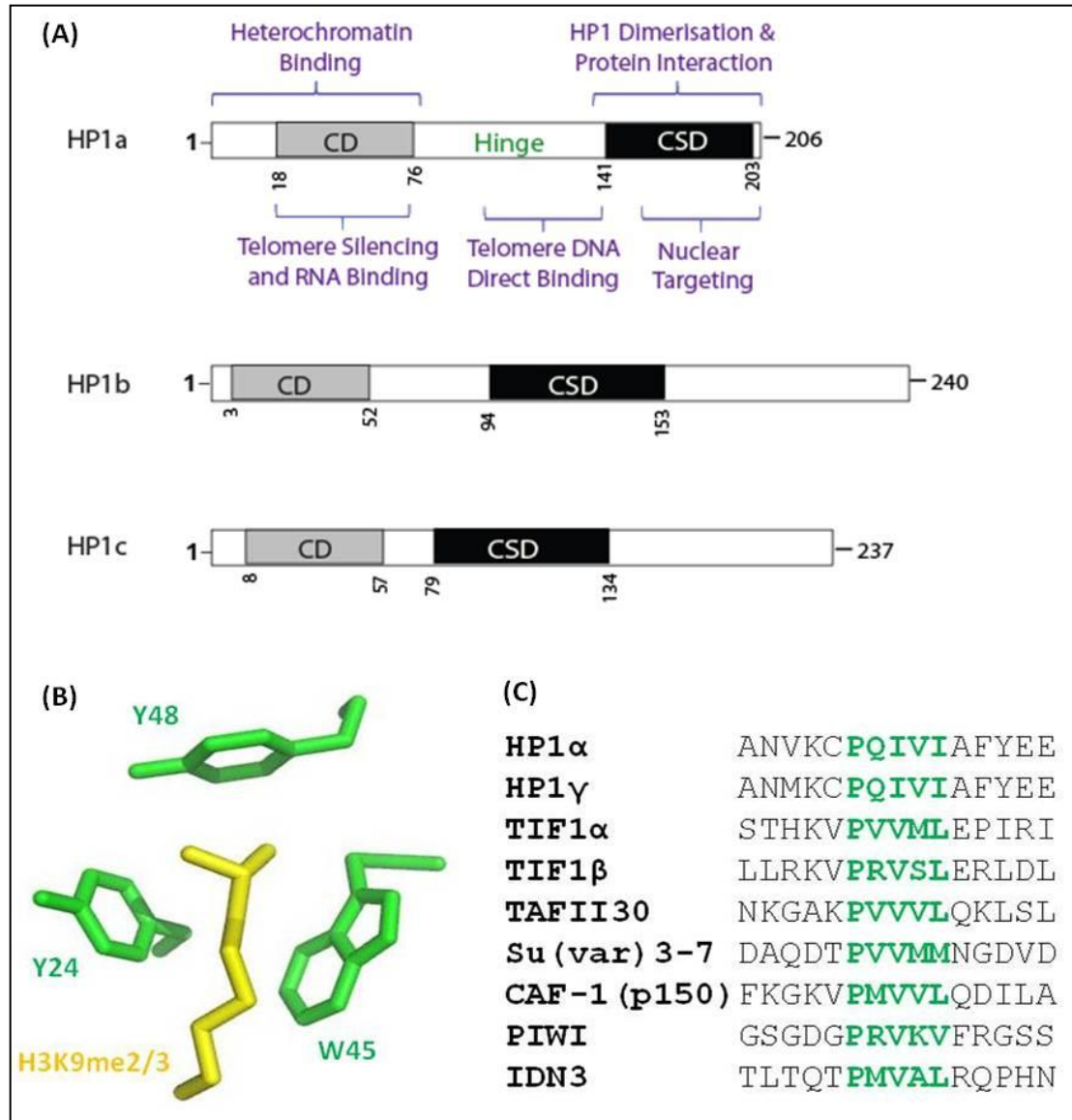
HP1 is a small chromosomal protein first discovered in *Drosophila melanogaster* as a major component of heterochromatin. Mutation of the fly gene *Su(var)2-5*, which encodes the HP1 protein, suppressed the silencing effect of heterochromatin in PEV. HP1 was therefore found to function as a dosage-dependent modifier of PEV in heterochromatic domains (Eissenberg et al., 1990). HP1 is phylogenetically highly conserved and present in almost all eukaryotes with the exception of budding yeast (*Saccharomyces cerevisiae*), but not prokaryotes. Generally, the HP1 family of proteins act in the repression of gene activity via the establishment and maintenance of higher-order chromatin structures.

Multiple HP1 members can exist within the same species, for example there are three distinct isoforms in the *Drosophila* and mammalian HP1 families named HP1a,b,c and HP1 $\alpha,\beta,\gamma$  respectively (Figure 1.8A). HP1 proteins possess a characteristic domain organisation containing two highly conserved domains: an N-terminal CD and a C-terminal chromoshadow domain (CSD). These two domains are separated by a variable linker or hinge region containing a nuclear localisation sequence (NLS). The CD is a chromatin binding module that specifically recognises and binds to di- and tri-methylated lysine 9 of histone H3. The CD binds the methylammonium group of K9 via a cage of three aromatic residues that becomes ordered on binding. Other specific sites of the HP1 CD make contact with residues surrounding the K9 site of histone H3 and it is these interactions that appear to be the key determinant of recognition specificity of this reader module (Nielsen et al., 2002). The interaction between the CD and the methyl K9 histone H3 mark is

important for the recruitment of HP1 to heterochromatic regions of the genome and its overall gene-silencing function (Kwon and Workman, 2011).

The structure and globular conformation of the HP1 CSD is similar to that of the CD however, although the CD remains monomeric in solution the CSD has been shown to readily dimerise under the same conditions (Cowieson et al., 2000). Upon HP1 CSD dimerisation, a prominent non-polar groove is generated at the dimer interface that can accommodate HP1-interacting proteins containing the extended pentapeptide binding motif PxV $\psi\psi$  (where  $\psi$  represents a hydrophobic residue) (Cowieson et al., 2000). The CSD has been implicated in a wide array of protein-protein interactions with factors that include transcriptional regulators, chromatin modifiers, replication and cell-cycle-related factors and nuclear architecture proteins. It is suggested that the CSD may be responsible for the targeting of HP1 to specific binding sites within chromatin (Kwon and Workman, 2011).

The linker region separating the highly conserved CD and CSD is less conserved and contains the most variable primary sequence between HP1 proteins. Structurally, the hinge region has been proposed to be flexible and exposed to the surface (Singh and Georgatos, 2002) and is thereby highly susceptible to post-translational modifications. Modifications have been shown to impact on the localisation, interactions and functions of HP1, highlighting the linker region as a possible central control region in the regulation of HP1 proteins (Lomberk et al., 2006).



**Figure 1.13. Schematic representation of HP1 paralogs in *Drosophila*.** (A) The HP1 protein contains an N-terminal chromodomain (CD) and a C-terminal chromoshadow domain (CSD), separated by a hinge region. The functional roles of the different domains are described. (B) Structural recognition of chromatin by the CD is depicted. The binding pocket for the *N*-methyl group is provided by three aromatic side chains.(PDB 1KNE). (C) The CSD is required for HP1 dimerisation and interaction with many proteins that possess a conserved PxV $\psi\psi$  penta-peptide signature. Sequence alignments illustrate acknowledged HP1 binding proteins containing this consensus.

The localisation and function of HP1 proteins at heterochromatic regions has been well documented. However, emerging studies show that they localise not only to heterochromatic domains but also to euchromatic domains (Lomberk et al., 2006). The localisation pattern appears to be isoform-specific, with mammalian HP1 $\alpha$  and  $\beta$  being predominantly heterochromatic and HP1 $\gamma$  being found in euchromatin (Fanti et al., 2003). The localisation of HP1 is regulated by a number of distinct factors. Firstly, at an epigenetic level, the major mechanism localising HP1 within centric heterochromatin is via its CD interaction with methylated lysine 9 of histone H3. Recently, HP1 $\alpha$  has also been shown to specifically bind to the H3Y41 region via its CSD and JAK2-mediated phosphorylation of this residue dissociates HP1 $\alpha$  from chromatin (Dawson et al., 2009). The revelation of novel HP1 recruitment factors demonstrates the complexity involved in targeting HP1 to distinct regions in different contexts. Alternative mechanisms defining the subnuclear localisation of HP1 involve post-translational modification of the HP1 protein itself, interactions between the CSD and other protein factors and at the genetic level, direct binding of HP1 to DNA (Badugu et al., 2003) and RNA transcripts (Piacentini et al., 2009).

The most common of HP1 functions is in the formation of heterochromatin. The underlying basis for heterochromatin formation and gene silencing revolves around interactions between the HMTase SU(VAR)3-9, HP1 and the H3K9me3 modification. According to the model, SU(VAR)39 methylates histone H3 at lysine 9, creating a binding site for HP1 CD whilst interacting itself with the HP1 CSD. This three component complex then forms and allows the spreading of higher order chromatin states (Fanti and Pimpinelli, 2008). A recent study implicates HP1 at the hub of the spread and assembly of heterochromatin containing H3K9 methylation.

Canzio *et al.*, have shown that the *S. pombe* HP1 homolog, Swi6, recognises the H3K9me3 mark at the interface of two CD. This interaction causes Swi6 to tetramerise on a nucleosome, generating two CD sticky ends which can bridge nearby methylated nucleosomes (Canzio *et al.*, 2011). HP1 CD oligomerisation therefore enhances the spread of heterochromatin *in vivo*.

In addition to functioning in the formation of heterochromatin and heterochromatic gene silencing, HP1 functions in the regulation of gene expression at multiple euchromatic sites. In general, HP1 proteins have been shown to decrease RNAP occupancy at gene promoters by recruiting silencing factors and creating less accessible localised chromatin structures. More specifically, HP1 $\gamma$  was found to associate with the coding regions of transcriptionally active genes. Tethering of HP1 $\gamma$  upstream of a gene promoter prompted silencing concurrent with localised chromatin condensation and an increase in H3K9 methylation (Maison and Almouzni, 2004). In depth analysis of such events implicated HP1-mediated silencing through the targeting of RNAP II coactivator complexes. HP1 was found to inhibit the assembly of the PIC *in vitro*, primarily blocking the addition of key subunits of the TFIID and Mediator complexes (Smallwood *et al.*, 2008).

A growing body of evidence is also implicating HP1 as a positive regulator of gene expression at a subset of euchromatic loci. High resolution mapping experiments have shown that HP1 is associated with transcriptionally active chromatin in *Drosophila* (de Wit *et al.*, 2007). Among its numerous euchromatic binding sites, detailed mapping has shown its enrichment at developmentally regulated chromosome puffs in polytene chromosomes; structures that represent sites of intense gene activity. HP1 binds to heat-shocked-induced puffs after induction and



is positively involved in *Hsp70* gene activity (Piacentini et al., 2003). Furthermore, HP1c has been shown to interact with the histone chaperone complex FACT. HP1 guides the recruitment of FACT to active genes and links FACT to active forms of RNAP II. In the absence of HP1c the recruitment of FACT into heat shock loci is diminished, which subsequently causes a defect in heat shock gene expression (Kwon et al., 2010). Collectively, these studies implicate a role for HP1 in positive gene expression; however how HP1 proteins mediate this process still remains ambiguous.

A number of recent studies highlight an emerging role for HP1 proteins in the DNA damage response (DDR). Several of these studies have established that heterochromatin is inhibitory to the repair process and that reducing levels of HP1 results in the relaxation of heterochromatin, thus bypassing the requirement for ATM signalling and enhancing DNA repair (Ayoub et al., 2008). Ayoub *et al.* showed that DNA-damage induced phosphorylation of the CD located T51 residue in HP1 $\beta$  triggers its dissociation from chromatin by inhibiting its interaction with H3K9me3. Following this transient dispersal of HP1 $\beta$  from the H3K9me3 modification, HP1 $\beta$  reaccumulates at damaged sites to restore the levels of bound HP1 to heterochromatic regions.

In contrast, more recent studies reveal that HP1 proteins may be involved in facilitating the DNA repair pathways in response to different genotoxic stresses. All three isoforms of HP1 are recruited to UV-induced DNA lesions in both human and mouse cells (Dinant and Luijsterburg, 2009). Despite there being some HP1 dissociation from H3K9me3 at damaged sites, recent results indicate a net accumulation of HP1 through its CSD at damaged chromatin. Moreover, the importance of HP1 in the response to UV damage is emphasised by the dramatic

sensitivity of nematodes, to UV, lacking functional HP1 proteins (Luijsterburg et al., 2009).

### 1.14. Project Aims and Objectives

MOZ is an essential coactivator for genes that are involved in haematopoietic and developmental processes. However, the understanding of the mechanisms involved in MOZ function is vague. The dysregulation of MOZ is intimately associated with AML, where the oncogenic fusion proteins responsible for AML include the entire N terminus of MOZ. Within this segment are three functional domains, the H1/5, double PHD finger (DPF) and MYST domain, of which the latter is responsible for the proteins biochemical acetyltransferase function. The function of the MYST domain has been studied in some depth, however an anomaly still remains as to the specific target(s) of MOZ-acetylation; a result of conflicting *in vitro* and *in vivo* data. In comparison, little is currently known about the functions of the other two regions that are present in the fusion protein. Defining the role of the MOZ DPF module will shed light on its possible role in target specificity, how MOZ functions as a modular protein and consequently, give valuable insight into the mechanisms of carcinogenesis associated with MOZ fusion proteins.

At the outset of this project it was not known if MOZ could bind histone proteins. Thereby, the primary aim of this project was to establish the function of the MOZ DPF by carrying out a detailed biochemical characterisation of the interaction between the DPF and chromatin. Secondly, we aimed to investigate whether the DPF domain influences other biochemical functions of the MOZ protein, including its

HAT activity, transcriptional coactivator capabilities and subcellular localisation. The final objective was to explore a possible interaction of the MOZ protein with transcriptional repression regulators and to identify the functional consequence of such interactions.

## **Chapter Two: Materials and Methods**

### 2.1. Sources of Materials

#### 2.1.1. General suppliers

All general laboratory chemicals were of analytical grade and supplied by Fisher Chemicals or Sigma Aldrich unless otherwise stated. Phosphate Buffer Saline (PBS) was prepared using PBS tablets supplied by OXOID Ltd. Double deionised water was used to make all solutions, unless RNase free water was required and then deionised water was purified by passage through a Purite Neptune water purification system. The pH of solutions was measured using a pH meter (Jenway 3510) against solutions of a known pH. Where appropriate, sterilisation was achieved by filtration through 0.2 micron filter or by autoclaving for 20 minutes at 120 °C.

#### 2.1.2. Bacterial reagents

The bacterial growth medium components, tryptone and yeast extract were purchased from OXOID Ltd. *Escherichia. coli* strain DH5 $\alpha$  (Hanahan, 1983) was purchased from Stratagene.

#### 2.1.3. Molecular biology reagents

All restriction enzymes and their respective buffers were purchased from Roche Applied Science or New England Biolabs (NEB). Polymerase enzymes and their respective buffers and supplements were purchased from Finnzymes or Invitrogen Life Sciences. Molecular size markers and dATP, dCTP, dGTP and dTTP were purchased from Invitrogen Life Sciences. Gel extraction kits were obtained

from York Biosciences, Nucleospin<sup>®</sup> plasmid kits from Macherey-Nagel and maxi-DNA kits from QIAGEN Ltd.

### 2.1.4. Tissue culture reagents

All tissue culture supplements were purchased from Lonza Biowhittaker<sup>®</sup> and plastic-ware was obtained from Helena Biosciences.

### 2.1.5. Biochemical reagents

30 % (w/v) acrylamide mix was purchased from National Diagnostics. SDS-PAGE electrophoresis apparatus, Precision Protein Standards, wet transfer apparatus and Nitrocellulose transfer membrane were all purchased from Bio-rad. Complete Protease Inhibitor cocktail tablets were obtained from Roche. All primary antibodies, Horseradish peroxidase (HRP) conjugated secondary antibodies, Protein A or G sepharose beads were purchased from Abcam, Sigma-Aldrich, Roche and Millipore. Radioactive isotopes and Amplify solution were obtained from GE Healthcare. Biotinylated histone peptides were purchased from Millipore or custom made by Peptide Protein Research and fluorescein-labelled histone peptides were also custom produced by Peptide Protein Research.

### 2.2. Bacterial preparation and culture

#### 2.2.1. Composition of solutions and media used in bacterial procedures

**Luria-Bertani (LB) medium:** 1 % (w/v) tryptone, 0.5 % (w/v) yeast extract 1 % (w/v) NaCl (pH adjusted to 7.0). The medium was solidified by the addition of 2 % (w/v) bacteriological agar where required.

**Solid Growth Medium:** For LB agar plates 1 L of LB medium was supplemented with 15 g bacto-agar prior to sterilisation. The appropriate antibiotic was added to the molten agar medium and 20-25 ml was poured into each petri-dish and allowed to solidify at room temperature. Plates were stored at 4 °C for up to 1 month.

**Ampicillin:** 1000x stock solution: 100 mg/ml in 50 % ethanol.

**Chloramphenicol:** 1000x stock solution: 34 mg in 100 % ethanol.

**Kanamycin:** 1000x stock solution: 50 mg/ml in H<sub>2</sub>O.

#### 2.2.2. Culture of *Escherichia coli* strains

The bacterial strain *E. coli* DH5 $\alpha$  was used for all DNA manipulations. Alternatively, various BL21 strains including the standard BL21(DE3), B834, Codon+ and Rosetta were used in optimised protein expression protocols. Single colonies of bacteria were obtained by streaking liquid culture onto LB agar plates containing the appropriate antibiotic and grown at 37 °C. Liquid cultures were grown by inoculating LB (plus antibiotic) medium with a single bacterial colony. Cultures were then incubated in universal tubes (Bibby Sterilin, Staffordshire, UK) or conical flasks at 37 °C with shaking at 220 rpm.

### 2.2.3. Preparation of *Escherichia coli* competent cells

*E. coli* DH5 $\alpha$  cells were made competent for DNA transformation using the calcium chloride-magnesium chloride method. A single colony was used to inoculate a 2.5 ml LB medium and grown overnight at 37 °C. The overnight culture was diluted 1:100 in LB medium and grown at 37 °C to an Optical Density (OD<sub>600 nm</sub>) of 0.3. The cells were chilled on ice for 20 minutes and consequently harvested by centrifugation at 4 °C, 3500 rpm for 5 minutes. On ice, cell pellets were resuspended in 20 ml of ice cold 100 mM MgCl<sub>2</sub> and incubated on ice for 5 minutes. After a second centrifugation, cells were resuspended in 10 ml ice cold 100 mM CaCl<sub>2</sub> and incubated on ice for 30 minutes. The cells were harvested by centrifugation and finally resuspended in 5 ml CaCl<sub>2</sub> solution (70 mM CaCl<sub>2</sub>, 15 % glycerol). Microfuge tubes were placed in liquid nitrogen to allow snap freezing of 200  $\mu$ l cell suspension aliquots. Competent cells were stored at -80 °C. The transformation efficiency of the cells was checked using pcDNA3.1+ vector and the transformation serially diluted to allow the determination of the number of transformants: 1  $\mu$ g DNA should give  $\sim 1 \times 10^6$  transformants.

### 2.2.4. Transformation into *Escherichia coli* cells

DNA products were transformed into competent DH5 $\alpha$  or BL21(DE3) *E. coli* strains. The competent host cells were thawed on ice. For each of the transformations 0.5  $\mu$ g of plasmid DNA was mixed gently with a 50  $\mu$ l cell suspension volume and incubated on ice for 30 minutes. The cells were heat shocked at 42 °C for 2 minutes before returning to ice for 2 minutes. After the addition of 500  $\mu$ l of LB medium the



## Chapter Two: Materials and Methods

reactions were incubated at 37 °C for 1 hour. Transformations were spun at 13,000 rpm for 1 minute to pellet the cells and 350 µl of supernatant was discarded. Cells were resuspended in the remaining LB medium and spread onto LB agar plates containing the appropriate antibiotic (ampicillin/ chloramphenicol/ kanamycin). The plates were incubated overnight at 37 °C.

### 2.2.5. Long-term storage of bacterial cultures

To preserve transformed bacterial strains for long periods of time, glycerol stocks were prepared. Stocks consisted of a mixture of the bacteria in broth with sterile glycerol at an 80 % to 20 % ratio respectively. The contents were vortexed and stored at -80 °C.

## 2.3. Molecular biology techniques

### 2.3.1. Composition of solutions used in molecular biology methods

**Gel loading buffer (5x):** 15 g Ficoll400 dissolved in 10 ml 10 mM Tris-HCl-pH7.5 and 50 mM EDTA-pH8.0, containing 0.4 % orange G, 0.03 % bromophenol blue and 0.03 % xylene cyanol FF.

**Tris-borate-EDTA (TBE):** 40 mM Tris-base, 40 mM Boric Acid, 0.5 mM EDTA (pH8.0).

### 2.3.2. Small-scale plasmid purification from bacterial cells

*E. coli* cultures of 5 ml were grown overnight in an orbital shaker at 37 °C, 220 rpm. Cells were harvested by centrifugation and the plasmid DNA was recovered from the pellet using the NucleoSpin<sup>®</sup> Plasmid kit (Macherey-Nagel), according to manufacturer's instructions. DNA was eluted from the NucleoSpin column with the supplied Elution Buffer (EB).

### 2.3.3. Large-scale plasmid purification from bacterial cells

To obtain a large-scale preparation of plasmid DNA, 250 ml LB media with appropriate antibiotic was inoculated with 1/100 fold dilution of suspended cells and left to grow overnight at 37 °C, 220 rpm. To isolate plasmid DNA from these cells the QIAGEN Plasmid Maxi Kit protocol was used, involving alkaline lysis of the cells, followed by immobilisation of the plasmid DNA onto a QIAGEN anion-exchange resin. Elution and precipitation of the DNA was achieved in isopropanol, prior to being washed in 75 % ethanol and dried under vacuum. The resulting DNA pellet was resuspended in sterile water. Details of this protocol and the composition of buffers used can be found in the QIAGEN maxiprep handbook.

### 2.3.4. Phenol/chloroform extraction and ethanol precipitation of DNA

To separate nucleic acids from contaminating proteins, a phenol/chloroform extraction was carried out. An equal volume of phenol:chloroform:isoamyl alcohol (25:24:1 saturated with 10 mM Tris pH8.0, 1 mM EDTA) solution was added to the

DNA sample to extract protein contaminants. The sample was vigorously mixed until an emulsion was formed and spun at 13,000 rpm for 2 minutes in a desk-top centrifuge (Heraeus Biofuge *pico*). The aqueous layer was transferred to a new microfuge tube and a further round of phenol:chloroform extraction carried out until all protein was removed from the interphase. Following the preceding protocol, several rounds of chloroform extraction were used to remove residual phenol. The DNA was then collected by ethanol precipitation.

### 2.3.5. Ethanol precipitation

DNA was precipitated using 0.1 volumes of 3 M NaOAc (pH5.2) and 2 volumes of 100 % ethanol. The precipitate was recovered after vigorous mixing by centrifugation for 30 minutes at 13,000 rpm, 4 °C (ALC PK131R, rotor T528). To remove excess salt ions, the pellet was washed with 75 % ethanol and the centrifugation repeated. The pellet was dried under vacuum in a heated speed-vac (Savant DNA120 speed-vac) for 5 minutes before resuspension in sterile water.

### 2.3.6. Spectrophotometric quantification

The quantification and quality of DNA was determined by optical density measurement at 260 nm using the **NanoDrop**<sup>®</sup> ND-1000 UV-Vis Spectrophotometer. These parameters were also clarified by restriction digest (section 2.3.13) or by comparison of DNA samples with molecular weight markers of known concentration following separation by agarose gel electrophoresis (section 2.3.8). OD<sub>260</sub> nm

/OD<sub>280</sub> nm values were also calculated and were useful indicators of protein sample purity.

### 2.3.7. Sequencing of plasmid DNA

Plasmid DNA for sequencing was sent to Geneservice Ltd (Source Bioscience Ltd Nottingham, UK) and was sequenced using Big Dye V3.1 chemistry and an AB13730xl Automated Sequencer.

### 2.3.8. Agarose gel electrophoresis

DNA was fractionated according to size on neutral agarose gels of between 0.7 % and 2 % agarose (w/v) depending on the size of fragments being resolved. Gels containing 0.5 µg/ml ethidium bromide were set and run in 0.5x TBE buffer as described in standard protocols (Sambrook J, 1989). DNA samples were loaded in 5x gel-loading buffer, alongside Invitrogen standard 1kb DNA ladder (0.2 µg/µl). Bands were then visualised using an ultraviolet transilluminator (BioRad Gel Doc 2000) and BioRad Quantity One computer software. The approximate size (kb) and concentration (ng) of DNA fragments were determined by comparison with standard molecular weight markers.

### 2.3.9. Isolation and purification of DNA from agarose gels

DNA bands to be purified were visualized using an ultraviolet light box and excised from the gel after separation by electrophoresis. DNA was purified using the York Biosciences Gel/PCR purification kit according to manufacturer's instructions.

### 2.3.10. Oligonucleotides

Oligonucleotides were purchased from Sigma Genosys as lyophilised and desalted pellets. They were resuspended in sterile water to a 100  $\mu$ M concentration and stored at -20 °C. Oligonucleotides used in this work are listed in appendix tables A.1 and A.2.

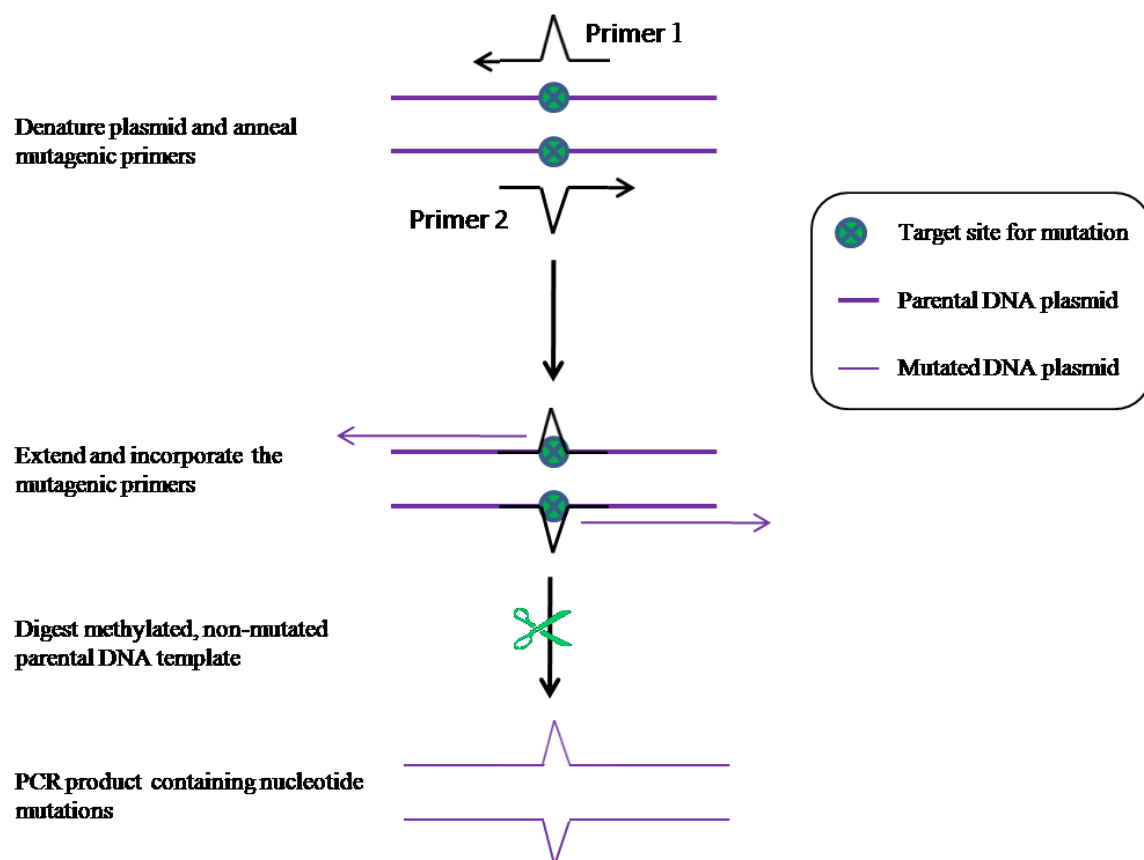
### 2.2.11. Polymerase chain reaction (PCR)

PCR was used to generate DNA fragments needed for the construction of recombinant plasmids. To synthesise DNA for cloning purposes, a DNA polymerase enzyme with proofreading 3' to 5' exonuclease activity, Phusion<sup>®</sup> High-Fidelity Polymerase (Finnzymes) was used. A typical reaction mixture consisted of 150 ng of DNA template and a final concentration of 1x Phusion HF reaction buffer, 0.5  $\mu$ M of both the forward and reverse PCR primers, 0.2 mM of each dNTP, 2 % DMSO, 1.25 units of Phusion<sup>®</sup> High-Fidelity DNA polymerase and made up to a final volume of 50 $\mu$ l with nuclease-free sterile water. Reactions underwent a 'hot start' PCR method in the Applied Biosystems 2720 thermal cycler. Samples were initially denatured by incubating at 98 °C for 30 seconds and then typical thermal cycling parameters were

as follows: denaturation, 98 °C for 10 seconds, annealing, 55 °C for 30s and extension, 72 °C for 30 seconds per kb of target DNA, for a total of 25-35 cycles. The temperature was maintained finally at 72 °C for 10 minutes. Once completed, 10 % of the PCR reaction was separated by agarose gel electrophoresis to verify that correctly sized DNA fragment had been generated.

### **2.3.12. Generation of mutations/deletions by recombinant PCR**

To mutate/delete specific nucleotides within a DNA sequence two oligonucleotides were designed (Appendix table A.2) and used as primers for PCR as shown in Figure 2.1. Primers 1 and 2 were complimentary, except for the nucleotide mutations and spanned the region to be mutated. PCR reactions were carried out to extend and incorporate the mutagenic primers, resulting in nicked circular strands of mutated plasmid DNA. PCR reactions were carried out as described in section 2.3.11 with the following amendments; 200 ng of DNA template was used and the thermal cycling extension time was increased to 20 minutes at 72 °C. PCR products were then treated with 0.5 µl Dpn1 restriction enzyme to digest the methylated, nonmutated parental DNA template for 2 hours at 37 °C. Control (no oligonucleotides used) and sample reactions were separated by agarose gel electrophoresis (section 2.3.8) to check for complete digestion of the parental template. The nicked, circular dsDNA was transformed into DH5 $\alpha$  (section 2.2.4) and the resulting single colonies used to prepare small-scale plasmid DNA preparations (section 2.3.2), which were then sent for sequencing (section 2.3.7) to check for the incorporation of the mutated nucleotide sequence.



**Figure 2.1. Generation of mutations by recombinant PCR.** Recombinant PCR can be used in order to create nucleotide mutations within a plasmid construct. Two complimentary primers are generated which contain the nucleotide mutation (illustrated by V /  $\Delta$ ) and utilised in the PCR reaction. During PCR, mutagenic primers are annealed to plasmid DNA and extended and incorporated by the action of the DNA polymerase, resulting in nicked circular DNA strands. This generates a full-length PCR product containing the mutation. Methylated, non-mutated parental DNA was then digested using the Dpn1 restriction enzyme and transformed into DH5 $\alpha$  *E. coli* in order to prepare the newly mutated plasmid DNA.

### 2.3.13. Restriction digests

Digests of plasmid DNA with restriction endonucleases were used to generate compatible ends for cloning and also to verify the insertion of a gene fragment into a newly created plasmid. The specific enzymes chosen were dictated by the primer set used in the cloning protocol and digests were carried out in the buffers and at the temperatures recommended by the manufacturer (Roche). Double digests were carried out together where the buffers and temperatures were compatible (both enzymatic activities > 75 %), otherwise the reactions were carried out sequentially.

Typically, digests were set up in a 20 µl reaction volume consisting of 1x SuRE/cut reaction buffer, 2 units restriction endonuclease (when a second enzyme was used, the total reaction volume was increased to prevent an excess glycerol concentration of 10 % that can interfere with the activity and specificity of endonuclease enzymes), template DNA (concentration depending on purpose) and sterile water up to a final volume of 20 µl. Digests were commonly carried out at 37 °C for 2 hours unless advised otherwise by the manufacturer.

### 2.3.14. Phosphatase treatment

Digested vectors were prepared for ligation by treatment with 1.0 µl of Thermosensitive Alkaline Phosphatase (Promega, 1 U/µl - TSAP) in 1x Promega Multicore™ buffer for 30 minutes at 37 °C. TSAP catalyses the removal of the terminal phosphate groups of DNA, thus preventing reannealing and self ligation of linearised cloning vector DNA. Reactions were incubated at 74 °C for 20 minutes to irreversibly inactivate TSAP.



### 2.3.15. Ligation reactions

Recombinant plasmids were generated by annealing cut fragments with T4 DNA ligase (Promega). Relative quantities of cut vector and digested inserts were determined by agarose gel electrophoresis (section 2.3.8) and spectrophotometric quantification (section 2.3.6). Ligation reactions were set up incorporating a vector:insert ratio of 1:2. Typically, 30 ng of vector DNA and 60 ng of insert DNA was made up to 8  $\mu$ l with sterile water and incubated at 50 °C for 5 minutes. T4 DNA ligase (1  $\mu$ l) and 1x T4 DNA ligase reaction buffer (1  $\mu$ l) was added and the ligations incubated for ~12 hours at 12 °C. Vector only control ligations were also set up as above except without the addition of insert DNA. The ligated DNA was transformed and expressed in DH5 $\alpha$  *E. coli* (section 2.2.4), plasmid preparations made (section 2.3.2) and verified for the correct integration of the gene fragment (2.3.13 & 2.3.16).

### 2.3.16. Colony PCR screening

Ligation colonies were picked from LB agar plates and used to inoculate 200  $\mu$ l of LB (+antibiotic) cultures. These cultures were grown at 37 °C for 2 hours or until the culture was turbid. The turbid culture was used as template, in place of DNA in proceeding PCR reactions. Taq<sup>TM</sup> (*Thermus aquaticus*) DNA polymerase (Invitrogen) was used to screen for bacterial cultures containing the appropriate DNA insert. Reactions consisted of the same reagents as previously stated with the following amendments: 1.25 units of Taq DNA polymerase and its corresponding buffer (1x), 1.5 mM MgCl<sub>2</sub> and 2.5  $\mu$ l of bacterial culture instead of DNA template. Positive clones were identified by running the entire PCR reaction on an agarose gel

(section 2.3.8). Correct insert integration was checked by diagnostic digests (section 2.3.13).

## 2.4 Cell culture

### 2.4.1. Maintenance of adherent cell lines

The human osteosarcoma, U2OS, and human embryonic kidney, HEK293 cell lines were routinely maintained in Dulbecco's Modified Eagle Medium (DMEM) containing phenol red and supplemented with 1 % (v/v) 0.2 M L-glutamine, 1 % Penicillin/Streptomycin antibiotics and 10 % (v/v) heat inactivated fetal calf serum (FCS). Cultures were grown in sterile plasticware at 37 °C in a humidified atmosphere containing 5 % CO<sub>2</sub>. Adherent cells were passaged at 90 % confluency for a maximum of 30 passages, at which point the cells were discarded. To split cells, DMEM was aspirated and the cells washed twice with sterile 1x PBS. 1 ml trypsin/EDTA was added (except to HEK293 cell line) and the cells incubated for 2 minutes at 37 °C, 5 % CO<sub>2</sub>. Trypsin was aspirated and the cells were resuspended and subsequently diluted in fresh DMEM in 10 cm<sup>2</sup> plates (for maintenance) or 6/12-well plates (for transient transfections).

### 2.4.2. Maintenance of suspension cell lines

The human promyelocytic leukaemia, HL-60 cell line were routinely maintained in Roswell Park Memorial Institute medium (RPMI) containing phenol red, supplemented with 1 % (v/v) 0.2 M L-glutamine, 1 % Penicillin/Streptomycin

antibiotics and 10 % (v/v) heat inactivated FCS. Cultures were grown in sterile plastic flasks at 37 °C in a humidified atmosphere containing 5 % CO<sub>2</sub>. Suspension cells were passaged every four days for a maximum of 30 passages, at which point the cells were discarded. To passage, cells were first counted. Cells were diluted in a 1:1 ratio and mixed thoroughly with trypan blue (Autogen Bioclear), loaded onto a hepatocytometer (Hawksley, BS.748) and counted under a light microscope (Nikon eclipse TS100). Cells were then diluted in fresh RPMI media to a concentration of 100,000 cells/ml in a sterile tissue culture flask and incubated at 37 °C, 5 % CO<sub>2</sub> (for maintenance). To harvest suspension cells for whole cell extract preparation, cells were collected at 1250 rpm for 10 minutes at room temperature and the RPMI media aspirated.

### 2.4.3. Cryopreservation of cell lines

Cells were incubated with trypsin for 1 minute and after removal were returned to the incubator for 5 minutes. Cells were then resuspended in 2 ml of appropriate media (for a 10cm<sup>2</sup> dish), counted (section 2.4.2) and diluted to the appropriate concentration. 400 µl (40 %) of cell suspension in complete media was transferred to labelled cryovials containing 100 µl (10 %) DMSO and a supplement of 500 µl (50 %) FCS. The cell suspension was then gently pipetted to ensure even mixing of DMSO, FCS and media. The vials were then stored at -20 °C until suspensions were frozen solid. The cryovials were then placed at -80 °C overnight, after which the vials were stored in liquid nitrogen.

### 2.4.4. Resuscitation of cell lines

To recover cells from liquid nitrogen, cells were thawed at 37 °C and transferred into 10 ml of pre-warmed supplemented media. The suspension was mixed thoroughly and cells harvested by centrifugation at 1250 rpm, 5 minutes to remove residual DMSO that could hinder growth. Cells were then resuspended into 10 ml of fresh media and incubated at 37 °C, 5 % CO<sub>2</sub> with humidity. A few hours post cell plating the media was changed to remove dead and unattached cells. After 24 hours the media was then changed one final time, to remove any dead cells. Cells were passaged at least three times (left approximately for 12 days) before experimental use.

### 2.4.5. Calcium phosphate-mediated transfection of adherent cells

Adherent cells were seeded at 20 % confluency in 10 cm plates 24 hours prior to transfection. The following day, DMEM was replaced with fresh media 3-4 hours prior to transfection. For transient transfections, 50-80 % subconfluent cells were transfected by using the calcium phosphate coprecipitation method. Cells were transfected with X µg of the appropriate plasmid DNA (quantity requires optimisation for each plasmid) and empty vector to standardise the quantity of DNA in each sample. To each sample 12.5 % (v/v) 2 M calcium chloride solution was added and made up to a final volume of 800 µl with sterile nuclease-free water pH7.0. This mixture was added dropwise to an equivalent volume of 2x HEPES Buffered Saline (HBS) solution whilst vortexing gently. After a 20 minute incubation at room temperature, 1.5 ml was added dropwise to the appropriate cell plate.

Following a 16 hour incubation at 37 °C, 5 % CO<sub>2</sub>, cells were washed thrice with 1x PBS and maintained in DMEM for a further 24 hours. Cells were lysed at the appropriate time post transfection (24-48 hours).

### 2.5. Biochemical techniques

#### 2.5.1. Composition of solutions used in biochemical methods

**Lysis buffer:** 20 mM sodium phosphate (pH7), 250 mM NaCl, 30 mM sodium pyrophosphate, 0.1 % NP-40, 10 mM NaF, 0.1 mM Na<sub>3</sub>VO<sub>4</sub>, 1 mM PMSF, 1x Complete protease inhibitors.

**Cytosol/Nuclear (C/N) Fractionation Buffer A:** 10 mM HEPES pH7.9, 10 mM KCl, 0.1 mM EGTA, 1 mM PMSF, 1 mM DTT, 1x Complete protease inhibitors.

**C/N Fractionation Buffer B:** 20 mM HEPES pH7.9, 400 mM NaCl, 1 mM EGTA, 1 mM PMSF, 1 mM DTT, 1x Complete protease inhibitors, 10 % glycerol.

**RIPA Lysis Buffer:** 10 mM Tris-HCl pH8, 0.5 mM EGTA, 140 mM NaCl, 1 % Triton X-100, 0.1 % sodium deoxycholate, 0.3 % SDS, 1 mM PMSF, 1x Complete protease inhibitors.

**Chro-IP Lysis Buffer:** 5 mM Tris-HCl pH8, 85 mM KCl, 0.5 % NP-40, 1 mM PMSF, 1x Complete protease inhibitors.

**Wash Nuclei Buffer:** 10 mM Tris-HCl pH8, 0.5 mM EGTA, 200 mM NaCl, 1 mM PMSF, 1x Complete protease inhibitors.

**LOW Salt Buffer:** 150 mM NaCl, 20 mM Tris-HCl pH8, 1 % Triton X-100, 0.1 % SDS.

## Chapter Two: Materials and Methods

**HIGH Salt Buffer:** 500 mM NaCl, 20 mM Tris-HCl pH8, 1 % Triton X-100, 0.1 % SDS.

**LiCl Buffer:** 0.25 M LiCl, 10 mM Tris-HCl pH8, 1 % Sodium deoxycholate, 1 % NP-40.

**SDS-PAGE loading buffer (4x):** 62.5 mM Tris-HCl pH6.8, 40 % glycerol, 2 % SDS, 14.5 mM 2- $\beta$ -Mercaptoethanol and a trace of bromophenol blue.

**Tris-glycine-SDS PAGE running buffer (10x):** 250 mM Tris base, 2 M Glycine, 35 mM SDS pH8.3.

**Tris-glycine-SDS transfer buffer (standard):** 48 mM Tris base, 39 mM glycine, 0.037 % (w/v) SDS, 20 % (v/v) methanol.

**Tris-glycine-SDS transfer buffer (for high molecular weight protein transfer):** 48 mM Tris base, 380 mM glycine, 0.037 % (w/v) SDS, 20 % (v/v) methanol.

**Brilliant Blue<sup>®</sup> staining solution:** 0.25% brilliant blue<sup>®</sup> (w/v) in 45% methanol, 10% acetic acid.

**Destaining solution:** 45 % methanol, 10 % acetic acid.

**ECL developing solution:** 1 ml Tris-HCl pH8.5, 6  $\mu$ l H<sub>2</sub>O<sub>2</sub>, 22  $\mu$ l solution A (90 mM p-Coumari Acid in DMSO), 50  $\mu$ l solution B (250 mM Luminol in DMSO) to a final volume of 10 ml with ddH<sub>2</sub>O.

**Fix solution:** 45 % methanol, 10 % acetic acid.

**NTN buffer:** 20 mM Tris-HCl pH8.0, 100-500 mM NaCl, 0.5 % NP-40, 1x Complete Protease Inhibitors (Roche).

**Glutathione Elution Buffer:** Freshly prepared, 30 mM reduced glutathione, 50 mM Tris-HCl pH9, 150 mM NaCl, 0.1 % Triton X-100.

## Chapter Two: Materials and Methods

**Binding buffer:**, 50 mM Tris-HCl pH7.5, 300 mM NaCl, 0.1 % NP-40, 1 mM PMSF  
1x Complete Protease Inhibitors.

**Blocking solution (for western blotting):** 5 % (w/v) milk powder in 1x PBS.

**Stripping buffer (for western blotting):** 62.5 mM Tris-HCl pH6.8, 2 % SDS, 100 mM 2- $\beta$ -Mercaptoethanol.

**10x HBS solution:** 1.37 M NaCl, 50 mM KCl, 8 mM Na<sub>2</sub>HPO<sub>4</sub>, 55 mM D-glucose,  
210 mM HEPES.

**Mounting media:** 90 % Glycerol, 10% PBS.

**2x HAT Buffer:** 100 mM Tris-HCl pH8.0, 20 % glycerol, 20 mM butyric acid, 2 mM PMSF, 2 mM DTT.

### 2.5.2. Preparation of whole-cell extracts

Cells were harvested by aspirating the DMEM media and scraping the cell layer off in 1 ml sterile 1xPBS. Cells were collected by centrifugation at 13,000 rpm, the supernatant removed and washed twice in PBS. Cells were resuspended in cold lysis buffer (2x cell compact volume) and incubated for 30 minutes, 4 °C with rotation. The cells were then snap frozen in liquid nitrogen, thawed in cold water and run over an eppendorf rack 10 times to ensure efficient lysis. Samples were again snap frozen in liquid nitrogen and thawed in water a further two times. If samples were viscous an optional sonication step was included (3x 20 sec on/30 sec off). Cellular debris was pelleted by centrifugation at 12,000 rpm at 4 °C for 30 minutes. The cell-free extract (supernatant) was transferred to a fresh tube and stored at -80 °C.

### 2.5.3. Preparation of cytoplasmic and nuclear extracts

Cells were harvested by aspirating the DMEM media and scraping the cell layer off in 1 ml sterile 1x PBS. Cells were collected by centrifugation at 13,000 rpm, the supernatant removed and washed twice in PBS. The cell pellet was resuspended in ice-cold C/N buffer A (approximately 3x cell compact volume) and incubated on ice for 10 minutes. A 0.5% final concentration of NP-40 was added to each sample, vortexed for 20 s, ice for 5 minutes and finally vortexed for a further 20 s. Samples were then spun at 13,000 rpm for 30 s and the supernatant (cytosolic fraction) was transferred to a fresh eppendorf tube. The resulting pellet was resuspended in C/N Buffer B and incubated on ice for 40 minutes. During this time period samples were vortexed for 15 s every 5 minutes. Again, samples were spun at 13,000 rpm for 20 minutes at 4 °C and the supernatant (nuclear fraction, nucleic acid-binding fraction) transferred into an ice-cold eppendorf tube. In order to extract any remaining insoluble proteins from the cellular debris, the pellet was finally resuspended in RIPA buffer. Samples were sonicated for 8 minutes 15 s on/ 15 s off and the cellular debris pelleted by centrifugation at 13,000 rpm, 20 minutes, 4 °C. The supernatant (nuclear fraction, insoluble nuclear protein fraction) was transferred to a fresh tube and all fractions stored at -80 °C.

### 2.5.4. Protein concentration determination

Protein concentrations were determined using the Bradford assay method and the Bio-Rad dye concentrate according to manufacturer's instructions. Typically, reactions were set up consisting of 2 µl of sample and 1 ml 20 % dye in ddH<sub>2</sub>O. The



samples were mixed thoroughly and left to stand at RT for 20 minutes. Absorbancies were read at 595 nm and the concentration calculated relative to values of diluted Bovine Serum Albumin (BSA) standards of known concentrations.

### **2.5.5. Sodium-dodecyl-sulphate polyacrylamide gel electrophoresis (SDS-PAGE)**

Cell extracts, Immunoprecipitations (IP), purified and *in vitro* translated proteins were analysed by one-dimensional polyacrylamide gel electrophoresis. SDS-PAGE gels were prepared with the appropriate percentage resolving and stacking gel (table 2.1), depending on the molecular weight of the protein being analysed. Protein extracts were denatured by boiling for 5 mins in 4x SDS-PAGE loading buffer prior to being resolved on an 8-15 % SDS-polyacrylamide gel. Gels were electrophoresised using a mini-gel system (Protean II, Bio-Rad) in 1x running buffer at a constant voltage.

### **2.5.6. Staining of SDS-polyacrylamide gels**

Gels were incubated in 0.25 % Brilliant blue stain for 30 minutes, followed by two rounds of 30 minute incubation in destain solution. To completely destain the gel, a final overnight incubation in distilled water was necessary. In order to preserve the stained gels, they were placed on filter paper, covered with Saran wrap and dried (80 °C, 1 hr) under vacuum (Biorad Hydrotech Vacuum Pump and Gel Dryer).

	Resolving Gel			Stacking Gel
% Acrylamide	8	12	15	
ddH <sub>2</sub> O (ml)	2.3	1.6	1.1	1.4
30% acrylamide mix (ml)	1.3	2.0	2.5	0.33
1.5 M Tris (pH 8.8) (ml)	1.3	1.3	1.3	-
1.0 M Tris (pH 6.8) (ml)	-	-	-	0.25
10 % (w/v) SDS (ml)	0.05	0.05	0.05	0.02
10 % (w/v) ammonium persulphate (ml)	0.05	0.05	0.05	0.02
TEMED (ml)	0.003	0.002	0.002	0.002

**Table 2.1. Solution compositions for resolving and stacking gels for SDS-polyacrylamide gel electrophoresis.** The volumes of the components stated are for the production of one 5 ml resolving gel and one 2 ml stacking gel.

### 2.5.7. Western blotting and immunodetection

Proteins separated on SDS-PAGE gels for western blotting were transferred to nitrocellulose membranes by electrophoresis in 1x transfer buffer using wet transfer apparatus. Membranes were stained in 1 % Ponceau S solution and then washed with sterile water to check the transfer efficiency. Ponceau S stained membranes were then photographed using the imager.

Nitrocellulose membranes containing immobilised proteins were incubated in blocking buffer for 1 hour at room temperature to block non-specific antibody

binding sites. Membranes were consequently incubated with primary antibody at 4°C overnight. Primary antibodies were prepared in blocking buffer at dilutions indicated in appendix table A.5. The nitrocellulose membrane was washed three times (3 x 5 min) in 1x PBS containing 0.1 % Tween 20, before the addition of the appropriate secondary antibody for 90 minutes at room temperature. HRP-conjugated secondary antibodies were diluted 1:3000 in blocking buffer. Excess secondary antibody was removed by three more washes (3 x 5 min) in 1x PBS containing 0.1 % Tween 20. Protein:antibody complexes were detected by ECL chemiluminescence. Membranes were covered for 1 minute with 5ml of ECL developing solution, excess reagent was drained off and the membrane wrapped in saran wrap. The chemiluminescent signal was then read by the Luminescent image analyser (Fujifilm LAS-4000). For re-use with other primary antibodies, membranes were stripped by incubation in stripping buffer for 30 minutes at 55 °C with occasional agitation. The membranes were then thoroughly washed and incubated in blocking solution as above and re-probed with a different primary antibody.

### **2.5.8. *In vitro* transcription and translation (IVT) of expression vector-encoded cDNA**

cDNAs cloned into expression vectors under the control of the T7 or SP6 promoter were transcribed and translated *in vitro* in a coupled reaction using the TNT<sup>®</sup> Coupled Reticulolysate System (Promega). Briefly, 1-2 µg of plasmid DNA (containing a T7/SP6 promoter) was mixed with 10 µl rabbit reticulolysate, 0.8 µl TNT reaction buffer, 0.4 µl T7/SP6 RNA polymerase, 0.4 µl amino acid mix (minus

Methionine), [35S]- Methionine (500  $\mu$ Ci/ 18.5 MBq, MP Biomedicals), 0.4  $\mu$ l RNasin Ribonuclease inhibitor (Promega) and nuclease-free water to 20  $\mu$ l. The mixture was incubated for 90 minutes at 30 °C. 1  $\mu$ l of each sample was then loaded onto an SDS-PAGE gel (section 2.5.5) to determine the efficiency of translation. The remaining IVT sample was stored at -80 °C.

### 2.5.9. Fixing and amplifying radioactive gels

Radioactive gels were fixed in 10 % (v/v) glacial acetic acid, 45 % methanol (v/v) solution for 30 minutes at room temperature. The gels were then placed for a further 30 minutes in 15 ml Amplify solution (Amersham). The gels were then placed on Whatmann filter paper, covered in Saran wrap and dried under a vacuum at 80 °C for 1 hour.

### 2.5.10. Exposure of radioactive gels to X-ray film

Dried gels were exposed to Kodak Biomax film in a Kodak Intensifying Screen (low energy) at -80 °C for between 16 hours and 1 week depending on the intensity of the signal of the IVT sample. The following day, the film was developed in the Curix 60 AGFA developer.

### 2.5.11. Co-immunoprecipitation

Protein extracts (750  $\mu$ g) obtained from mock and co-transfected cells were diluted to 500  $\mu$ l with lysis buffer. In order to pre-clear and remove unspecific extract

binding proteins, 10  $\mu$ l of Protein G Agarose (Roche) was added to each sample and incubated for 1 hour 4 °C with rotation. Samples were spun at 2,000 rpm for 2 minutes to pellet Protein-G beads and the extracts transferred to a fresh tube. Next, 2  $\mu$ g of antibody was added and samples were incubated overnight at 4 °C with rotation (a no antibody control was also carried out with the co-transfected sample). Protein-G beads were pre-washed in lysis buffer, 20  $\mu$ l added to each IP and samples incubated for a further 4 hours at 4 °C with rotation. Immunoprecipitations were washed thrice in lysis buffer and subsequently subjected to SDS-PAGE (section 2.5.5) and immunodetection (section 2.5.7) alongside a 10 % input as a control.

### 2.5.12. Chromatin immunoprecipitation (ChIP) assays

Cells were counted using a hepatocytometer and aliquoted into  $10.5 \times 10^6$  cells/ml in 50 ml falcon tubes. Cells were harvested at 2,000 rpm, 5 minutes at RT and subsequently, washed twice in 10 ml 1xPBS. To crosslink cells, 1 % formaldehyde was added dropwise and incubated with agitation for 10 minutes. A final concentration of 0.125 M glycine was then added for 5 minutes to stop the crosslink reaction. Cells were pelleted by spinning at 2,000 rpm, 5 minutes at 4 °C and washed twice in ice cold 1x PBS. At this stage of the protocol crosslinked samples could be snap frozen in liquid nitrogen and stored at -80 °C until ready for use.

To prepare chromatin from cells, they were firstly resuspended in 1 ml Chro-IP buffer and incubated on ice for 15 minutes. Cells were harvested by spinning at 2,000 rpm, 5 minutes, 4 °C, resuspended in 1 ml wash nuclei buffer and left on ice for

## Chapter Two: Materials and Methods

5 minutes. After further spinning, the pelleted cells were resuspended in 600  $\mu$ l RIPA lysis buffer, incubated on ice for 15 minutes and sonicated (10s on/30s off) on ice for 4 minutes. After sonication, cells were spun down at 14,000 rpm, 10 minutes, 4 °C and the supernatant (chromatin) transferred to a new tube. It was desirable for the chromatin to be chopped into DNA fragment sizes of 200-300 bp, so in order to check the size of the fragmented chromatin the following steps were carried out. Aliquots of 50  $\mu$ l of chromatin were incubated with 50  $\mu$ l of ddH<sub>2</sub>O and NaCl (final concentration 0.3 M) at 65 °C overnight to reverse the crosslink. 20  $\mu$ g of RNase was then added per 100  $\mu$ l sample and incubated for 10 minutes 37 °C. DNA was then purified with a PCR clean up kit (Macherey Nagle), eluted in 30  $\mu$ l EB and run on a 0.8 % agarose gel to check fragment size.

Following chromatin isolation, an IP assay was carried out. 50  $\mu$ l of Protein-G beads were used per IP reaction and 20  $\mu$ l per IP for the preclearing step. Prior to IP, Protein-G beads were blocked with 1 mg/ml BSA and 0.25 mg/ml sheared salmon sperm DNA and incubated overnight at 4 °C with rotation. 25  $\mu$ g of chromatin was diluted in RIPA buffer and precleared with 20  $\mu$ l preblocked Protein-G beads for 2 hours, 4 °C. Beads were harvested, the supernatant transferred to a fresh tube and incubated overnight with 2  $\mu$ g of antibody per IP at 4 °C. Preblocked beads (50  $\mu$ l) were then incubated with the IP samples for 2 hours at 4 °C. Beads were consequently washed once in RIPA, twice in LOW salt, twice in HIGH salt, twice in LiCl and twice in TE buffers. Following the stringent washing steps, ChIP samples were analysed by SDS-PAGE (section 2.5.5) and immunodetection (section 2.5.7).

### 2.5.13. Small-scale bacterial expression of GST-tagged proteins

Freshly transformed colonies were inoculated into 2.5 ml of LB-Ampicillin media. After 2 hours growth at 37 °C, 1 ml was placed into 2 microcentrifuge tubes. As a control 25 µl of water was added to one and 25 µl of 20 mM IPTG inducer (final concentration 0.5 mM) to the other. The remaining 0.5 ml was kept at 4 °C for future inoculation of large-scale cultures. After a further 2 hours growth at 37 °C, cells were harvested by centrifugation at 13,000 rpm for 5 minutes and the pellets resuspended in 15 µl 4x SDS loading buffer. Samples were boiled for 5 minutes and loaded onto an SDS-PAGE gel. Induced cultures were run alongside the appropriate control culture, followed by brilliant blue staining to detect the presence of an induced band.

### 2.5.14. Large-scale bacterial expression of GST-tagged proteins

The remaining 0.5 ml from bacterial cultures that showed good protein induction was used to inoculate 50 ml LB(+antibiotic) medium and left to grow overnight at 37 °C with 220 rpm shaking. The following day 1 L LB(+antibiotic) media was inoculated with the overnight culture and grown at 37 °C until cells reached exponential phase ( $OD_{600}$  0.8-0.9). A sample was transferred to a microcentrifuge tube (uninduced) and IPTG added to a final concentration of 0.3 mM. The cultures were incubated for a further 3-16 hours at 37-20 °C (conditions optimised for each protein) and subsequently, a second sample was taken representing the induced fraction. Cells were harvested by centrifugation at 5,500 rpm for 20 minutes at 4 °C (Sorvall Evolution RC rotor SLA-1500) and the pellet was frozen until ready for purification.

### 2.5.15. Purification of GST-tagged proteins

Induced bacterial pellets from 1 L cultures were resuspended in NTN buffer and sonicated for 5 minutes (30 sec on/ 30 sec off, Jenway). Cell debris was cleared by centrifugation at 18,000 rpm for 1 hour at 4 °C (Sorvall Evolution RC rotor SS-34), the supernatant transferred to a 15 ml falcon and a 20 µl sample removed. 200 µl of glutathione sepharose beads (GE Healthcare) were washed 4 times with an equal volume of 0.5 % milk in NTN buffer. On the last wash, beads were collected by centrifugation at 2,000 rpm for 2 minutes and resuspended in the NTN/milk buffer. 200 µl of bead slurry was added to the cleared bacterial supernatant and placed on a rotating wheel overnight at 4 °C. The glutathione beads were collected by centrifugation and a sample of the aspirated supernatant kept for future analysis. The beads were washed 3 times with NTN (2x 300 mM, 1x 100 mM NaCl) and finally, resuspended in 1 ml of the buffer and stored at 4 °C. SDS-PAGE was used to establish the level of Glutathione-S-Transferase (GST)-fusion protein induction and to monitor the success of purification at different stages of the procedure.

### 2.5.16. Elution and preparation of GST-tagged protein

To elute the GST-fusion proteins off the glutathione sepharose, beads were collected by centrifugation at 2,000 rpm for 2 minutes and the supernatant removed. 600 µl glutathione buffer was added to beads and the samples were placed on the rotating wheel overnight at 4 °C. Samples were spun to collect the beads and the supernatant transferred to a fresh eppendorf tube. A second round of elution was



carried out as above for 4 hours to ensure the complete recovery of GST-tagged protein.

The sample next underwent desalting and buffer exchange to remove low molecular weight compounds such as glutathione. This procedure was carried out using PD-10 desalting columns (containing Sephadex<sup>TM</sup> G-25 medium, Amersham). The bottom cap of the column was cut off, excess fluid removed and then the column was secured in a clamp. The column was equilibrated with approximately 25 ml of coupling buffer which was allowed to pass through by gravity-flow. The flow-through was discarded and a total volume of 2.5 ml protein sample (in binding buffer) was added to the column. Once again the flow-through was discarded and finally the GST-tagged protein was eluted with 3.5 ml of binding buffer and the flow-through collected.

The GST-tagged protein was then concentrated to a total volume of 500  $\mu$ l in a Vivaspin 2 ml concentrator (Vivascience). Samples were spun at 3,000 rpm, 4 °C for 3-6 hours (eppendorf centrifuge) until the sample volume had reduced to 500  $\mu$ l. The concentration of the protein was determined by Bradford assay and SDS-PAGE analysis against BSA standards of known concentration. Protein was next distributed into 20  $\mu$ l aliquots and stored at -80 °C.

### **2.5.17. GST-pulldown of histone extracts**

Normalised levels of GST-tagged fusion protein were determined by SDS-PAGE analysis. The desired amount of fusion protein (approximately 1.5  $\mu$ g) was incubated with 10  $\mu$ g calf thymus histone extracts (Sigma), overnight at 4 °C with

rotation in 750 µl of NTN buffer. The beads were then washed three times in NTN buffer, dried under vacuum (Savant DNA120 Speedvac), boiled in 20 µl of SDS-PAGE sample buffer and run on an SDS-PAGE gel to check for protein-histone interaction.

### 2.5.18. GST-pulldown of mammalian expressed protein

Normalised levels of GST-tagged fusion protein were determined by SDS-PAGE analysis. The desired amount of fusion protein (approximately 1.5 µg) was incubated with IVT protein (section 2.5.8) or 500 µg of whole cell extract (section 2.5.2), overnight at 4 °C with rotation in 750 µl of NTN buffer or cold lysis buffer respectively. The beads were then washed three times in the appropriate buffer, dried under vacuum (Savant DNA120 Speedvac), boiled in 20 µl of SDS-PAGE sample buffer and run on an SDS-PAGE gel to check for protein-protein interaction. Results were visualised via autoradiography (section 2.5.9-10) or western blotting (section 2.5.7).

### 2.5.19. Peptide binding assay

Normalised levels of *in vitro* translated protein or 1.5 µg of purified GST-tagged fusion protein (molarity values: GST- 75 nM; GST-MOZ 1-321- 32 nM; GST-MOZ DPF- 49 nM; GST-MOZ PHD1- 58 nM; GST-MOZ PHD2- 59 nM) was incubated with approximately 1.5 µg or 735 nM of biotin conjugated histone peptide (Millipore/Peptide Protein Research) or no peptide as control. Reactions were

incubated overnight with rotation at 4 °C in 750 µl Binding buffer (buffer composition section 2.5.1., p125). For simultaneous binding assays and titrations involving two peptides, a combined total of 735 nM peptide was used. A slurry of 15 µl Dynabeads<sup>®</sup> (M-280 Streptavidin-Invitrogen) were prewashed twice in 1x PBS + 0.1 % Tween and added to each reaction. The incubation with Dynabeads<sup>®</sup> took place for 1 hour at 4 °C. Dynabeads<sup>®</sup> were harvested using a magnet and consequently washed (3x10 minutes) in binding buffer. Samples were then subjected to SDS-PAGE (section 2.5.5) and results either visualised by autoradiography (sections 2.5.9-2.5.10) for IVT radio-labelled proteins or by immunodetection (section 2.5.7) with an  $\alpha$ -GST antibody (Sigma).

### 2.5.20. Luciferase reporter assays

Unless otherwise stated reporter assays were carried out in U2OS cells. Adherent cells were seeded at  $3 \times 10^5$  cells/well in 12-well plates. Transfections were carried out using the calcium phosphate coprecipitation method (section 2.4.5). Per well, cells were transfected with 500-1000 ng of MOZ expression plasmid, 500 ng of reporter plasmid (pT109 - 3xAML1-luc) and 100 ng of  $\beta$ -galactosidase ( $\beta$ -GAL) internal control plasmid. The total amount of DNA per well was standardised with the addition of pcDNA3.1(+) empty vector plasmid DNA. Reporter assays were carried out in triplicate for each condition. Samples were briefly vortexed, added dropwise to 2x HBS and incubated for 20 minutes prior to the addition to cells. Cells were washed in sterile PBS and fresh media added 16 hours post transfection. The cells were then lysed 48 hours after transfection. Lysis was carried out according to manufacturer's

## Chapter Two: Materials and Methods

instructions (Applied Biosystems). Briefly, cells were washed twice in PBS and 100  $\mu$ l lysis solution added to each well. Plates were then placed at  $-80^{\circ}\text{C}$  to snap freeze the lysates and stored at this point until ready to conduct the reporter assay.

Cells were then assayed using the Dual-light® Luciferase and  $\beta$ -GAL reporter gene assay system (Applied Biosystems). Plates were thawed and the remaining cells attached to the plates were scraped into the lysis solution and transferred to sterile eppendorf tubes. Lysates were spun at 12,000 rpm,  $4^{\circ}\text{C}$ , 5 minutes to pellet any cellular debris. The reporter assays were carried out in white-walled 96 well plates (Nunc). Per well 5  $\mu$ l of protein lysates was added, followed by 12.5  $\mu$ l of Buffer A. The assay was carried out according to instructions from here on, utilising the provided Buffer B with the addition of 1:100 galactone and accelerator solution. Luminescence was detected on a Berthold Orion Microplate Luminometer.

To analyse results, each Luciferase reading was corrected to the  $\beta$ -GAL value. The  $\beta$ -GAL dependent luminescence served as both a transfection efficiency control and a loading control for the assay. Values were expressed as Relative Luciferase Units (RLU). MOZ dependent transactivation of the reporter constructs was then expressed as a fold change between cells expressing MOZ and those transfected with the appropriate empty vector. This was necessary as some cellular factors were able to stimulate a basal activity of the Luciferase reporter. Protein expression was confirmed for each of the reporter lysates by SDS-PAGE analysis and immunoblotting (sections 2.5.5, 2.5.7).

### 2.5.21. Indirect immunofluorescence

$3 \times 10^5$  cells were seeded onto 25 mm coverslips in 6 well tissue culture plates and when fully adherent, were transfected with the appropriate plasmid DNA via the calcium-phosphate mediated method (section 2.4.5). 48 hours post-transfection the following steps were performed at room temperature with the coverslip undisturbed in the well. Cells were gently washed twice with PBS and fixed by incubation for 10 minutes with 4 % (w/v) paraformaldehyde. Cells were washed four times in PBS to remove excess fixative and permeabilised by incubation for 2 minutes with 0.2 % (v/v) Triton-x-100 (in PBS). Cells were washed as above, and non-specific binding sites blocked by incubation in 3 % (w/v) BSA (in PBS) for 30 minutes. Primary antibodies were prepared in 3 % (w/v) BSA (in PBS) blocking reagent at dilutions specified in appendix table A.5 to a final volume of 50  $\mu$ l. To prevent the coverslip drying out, a large sheet of blue towel was soaked in water and placed in a sealed container. 50  $\mu$ l of primary antibody was pipetted onto parafilm and coverslips were inverted (cell side down) onto the sample and incubated for 90 minutes in the sealed container creating a semi-humidified environment. The coverslips were replaced into the 6 well dish and washed as above. Alexa-488 or Alexa-594-conjugated secondary antibodies were prepared in blocking reagent at a 1:500 dilution and 500  $\mu$ l used to cover the cells. Cells were incubated for 30 minutes and covered with an aluminium foil box. Control cells were incubated in solely secondary antibody to determine the level of background fluorescence. The cells were washed again in PBS and finally incubated with 1:10,000 diluted Hoechst stain for 10 minutes. After a final set of washes in PBS the coverslips were removed from the wells, drained of excess

solution and placed cell side down onto 10  $\mu$ l mounting media spotted onto a microscope slide. Coverslips were then secured to the slide by adding a thin layer of clear nail varnish around the perimeter and allowed to dry. Slides were stored at 4 °C, protected from light. Immunofluorescence images were captured using a Zeiss LSM510 Meta confocal microscope and images were processed in the LSM Zeiss Image Browser.

### 2.2.22. Fluorescence polarisation

Fluorescence polarisation assays were performed for MOZ DPF binding to N-terminally fluoroscein labelled peptides of unmodified and modified histone H3 (amino acids 1-21) and histone H4 (amino acids 2-24). Protein concentrations were determined by SDS-PAGE analysis relative to the BSA protein standard. Fluorescence polarisation assays were performed in black polypropylene 384-well plates (Fisher) with 1 nM fluoroscein-labelled peptide and varying concentrations of MOZ DPF protein in binding buffer. Measurements were obtained after a 20 minute incubation at 25 °C using a fluorescence polarisation reader (Perkin Elmer).

### 2.5.23. *In vitro* histone acetyltransferase (HAT) assay

Approximately 1.5  $\mu$ g of purified HAT protein (molarity values: GST-MOZ DPF, 905 nM; GST-MOZ MYST, 620 nM; GST-MOZ DPF+MYST, 395 nM) was mixed with 1  $\mu$ l of a histone mixture (10 mg/ml) or 0.75  $\mu$ g histone peptide, 0.5  $\mu$ l BSA (10 mg/ml), 20  $\mu$ l 2x HAT buffer and made up to a final 40  $\mu$ l volume with

## Chapter Two: Materials and Methods

double-distilled water. The following control reactions were set up; Acetyl-CoA only, no HAT, and HAT without histones, in order to obtain background readings. Samples were then incubated for 2 minutes at 30 °C before the addition of 1.8 µl radio-labelled acetyl-CoA [ $^3\text{H}$ ] (15.4µM). Samples were incubated for a further 30 minutes at 30 °C. For kinetic assays samples were taken after the specified time at 30 °C. Next, 10 µl triplicates were aliquoted into eppendorfs containing 10 µl 4x SDS GLB and boiled to stop the reaction for 2 minutes. Each triplicate was spotted onto a filter (Whatman, 21mm) and allowed to dry for 1 hour. Filters were washed twice in 50 mM sodium carbonate pH9.2 solution and dried for 20-30 minutes at 60 °C. Each filter was finally placed in a scintillation vial (Perkin Elmer) containing 10 ml scintillation fluid (Ecoscint<sup>TM</sup>, National Diagnostics) to incubate overnight at room temperature. The radioactive decay (DPM) of tritium was counted using a Wallac 1409 liquid scintillation counter.

The remaining 10 µl from the acetylation assay was loaded onto and run on a 15% SDS-polyacrylamide gel. The gels were placed firstly in fixative, then in amplify and finally dried for one hour under vacuum. Gels were exposed to Kodak Biomax film in a Kodak Intensifying Screen (low energy) at -80 °C for 48 hours and the following day, the film was developed in the Curix 60 AGFA developer.

## **Chapter Three: Results**

### **Recognition of histone PTMs by the double PHD finger domain of MOZ**



### 3.1. Introduction

The MOZ MYST domain has been reported to target both histone and non-histone protein acetylation (Champagne et al., 2001; Ullah et al., 2008). In addition, the MYST module has been shown to possess DNA targeting activities, whereby the MYST domain binds to DNA via TFIIIA-type zinc finger and helix-turn-helix DNA-binding motifs located at the N- and C-terminal regions of the module (Holbert et al., 2007). However, how MOZ is recruited or targeted to chromatin substrates in order to direct its HAT activity is still unknown.

Other modules within the MOZ protein have also been assigned functions including the C-terminal amino-acid rich regions, which are essential for its transcriptional coactivator properties (Champagne et al., 2001) and the N-terminal H1/5 domain, which has been implicated in nuclear localisation and shown to possess an essential function in transcriptional coactivation (Yoshida and Kitabayashi, 2008). However, MOZ also contains N-terminal tandem PHD fingers of which the functional importance of the DPF domain is yet to be established; although they appear to be important for leukaemogenicity of MOZ fusion proteins (Deguchi et al., 2003).

The dysregulation of MOZ is intimately associated with AML, whereby the oncogenic fusion proteins responsible for AML retain the entire N-terminal region of MOZ, encompassing the H1/5, DPF, MYST and basic domains (Katsumoto et al., 2008). Furthermore, PHD fingers such as that in the NUP98-JARID1A fusion oncoprotein was found to be essential for leukaemic transformation and malignancy (Wang et al., 2009). These studies reveal a direct cause-effect relationship between a

deregulated PHD finger and oncogenesis. Thus, knowledge of DPF function is essential for our complete understanding of the molecular mechanisms involved in MOZ function in both a normal and malignant context.

The effector-mediated concept proposes that histone PTMs are ‘read’ and interpreted by protein modules, facilitating downstream events via the recruitment of non-histone binding effector proteins (Seet et al., 2006). PHD fingers are motifs predominantly found in nuclear proteins that function in the regulation of chromatin and are now well-established as chromatin binding modules (Bienz, 2006). The PHD domain has been characterised as a functionally versatile epigenome reader that can recognise and interpret the methylated (Li et al., 2006), acetylated (Lange et al., 2008) or the unmodified states (Lan et al., 2007) of histone lysine residues. In addition to MOZ and MORF, within the genome only three other proteins harbour a double PHD finger domain and they are the members of the d4-protein family, DPF1/2/3. The DPF3b tandem PHD fingers have been structurally and functionally characterised as a chromatin binding module, whereby the two PHD fingers function as an integrated unit that defines the histone PTM recognition specificity (Zeng et al., 2010).

Given a wealth of evidence showing that the PHD motif functions as an epigenome reader, this chapter describes work conducted in order to investigate whether the MOZ DPF domain can function as a histone PTM sensor. Furthermore, numerous studies were carried out to extensively characterise the specific binding pattern of histone PTMs recognised and bound by the MOZ DPF.

### 3.2. Sequence alignments of MOZ and other PHD fingers

PHD domains have been extensively studied both biochemically and structurally, resulting in the identification of distinct molecular mechanisms utilised for histone PTM recognition. Many PHD domains appear to have evolved to detect the modification status of H3K4. As discussed in the introduction, PHD domains that bind to unmodified H3K4 have been shown to make close contacts with residues in the vicinity of lysine 4, including hydrogen bonding and electrostatic interactions. In contrast, H3K4me3 binding PHD domains recognise the methyl groups via an aromatic cage and the consequent formation of favourable electrostatic cation- $\pi$  interactions. The MOZ PHD finger sequences were therefore aligned with other single PHD domains to determine if any structural features would give a clue to MOZ DPF function (Figures 3.1-3.2). All multiple alignments were conducted using the *MultAlin*, Expasy based program (Multiple sequence alignment with hierarchical clustering. F. CORPET, 1988, Nucl. Acids Res., 16 (22), 10881-10890). Sequence alignments indicate that MOZ PHD2 is similar to many other PHD fingers, whereas PHD1 is quite different to PHD2 and has fewer and lower identity homologues within the human genome. When aligned with PHD fingers that recognise the H3K4me3 PTM, such as ING family proteins or BPTF, it is evident that sequence similarity is very low (Figure 3.1). Sequences share 14 % similarity, of which 11 % are the structural cysteine and histidine residues that make up the universal PHD finger motif. None of the aromatic residues that configure the H3K4me3 binding cage are conserved in MOZ PHD finger 1 or 2 (Figure 3.1- **X**). Using this alignment we could predict that neither of the PHD fingers in the MOZ DPF unit function as H3K4me3

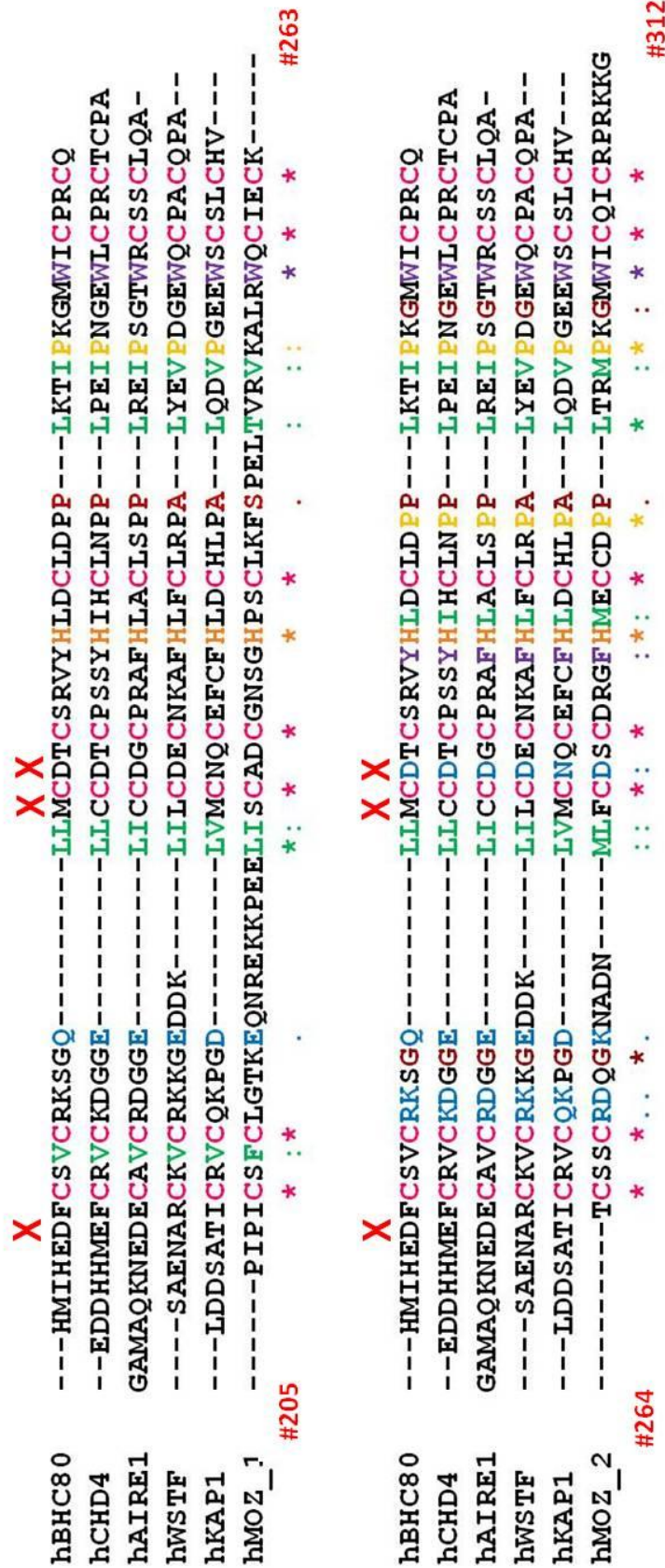
binding modules. Alignment with PHD fingers, such as BHC80 and AIRE1, that interact with the unmodified H3K4 residue show a greater sequence similarity compared to those detailed previously. Individually, the MOZ PHD1 shares a 27 % similarity with the other H3K4me0 binding PHD fingers, whereas MOZ PHD2 displays a greater similarity (37 %) (Figure 3.2). The majority of identical/similar residues are representative of the PHD motif (C4HC3) and those involved in binding and stabilising the N-terminus of the histone peptide (C-terminal similar residues form a hydrophobic pocket). The acidic residues involved in forming direct hydrogen bonds with the unmodified histone residue and mediating electrostatic interactions between backbone histone peptide residues are however, not conserved (Figure 3.2-**X**).

While most studies have focused on the functions of single PHD motifs, many proteins contain multiple copies of the PHD domain such as MLL, CHD4 and others. These can be clustered together or separated by other functional modules. MOZ and MORF are unusual in having tandem adjacent PHD fingers, however a protein NCBI BLAST search using MOZ DPF as an enquiry sequence presented the DPF2/3 proteins as high homology hits. Thus, this arrangement is shared with one other protein family, the d4-family members DPF1/2/3. As discussed earlier the DPF proteins contain two adjacent PHD fingers, although it should be noted that at the outset of this project the function of the DPF PHD fingers was unknown. The DPF tandem PHD fingers were later discovered to function as one unit in the molecular recognition of histone PTMs, whereby residues from both modules are required to establish the binding specificity of the entire domain (Zeng et al., 2010). Sequence alignments of human DPF proteins are shown in Figure 3.3. Alignment of the human

MOZ and DPF tandem PHD fingers highlights a high degree of conservation between the two primary sequences (Figure 3.3). DPF shares 53% identity and 72% similarity with MOZ DPF, whereby conservation specifically includes the majority of residues involved in the recognition of the H3 peptide N-terminus and acetylated H3K14. DPF3 is associated with the BAF chromatin remodelling complex and was identified as a novel reader of histone H3/ H4 acetylation and methylation. DPF3 represents the first PHD domain that recognises acetylated lysine residues, a property originally assigned to the bromodomain only (Lange et al., 2008).

[illegible]

**Figure 3.1. Sequence alignment of PHD fingers that recognise histone H3 in a H3K4 methylation dependent manner.** Identical residues are marked with an asterix (\*), strongly similar residues denoted with a colon (:) and weakly similar residues denoted with a dot (.). Zinc coordination cysteine and histidine residues are coloured in pink and orange respectively. Hydrophobic residues are coloured green, polar/charged residues in blue and aromatic residues in purple. MOZ amino acid residue numbers are denoted in red. **X** represents the aromatic/hydrophobic residues important for the site-specific recognition of H3K4me3. Multiple alignments were conducted using the *MultAlin*, ExPASy based program.



**Figure 3.2. Sequence alignment of PHD fingers that bind unmodified histone H3.** Identical residues are marked with an asterisk (\*), strongly similar residues denoted with a colon (:), and weakly similar residues denoted with a dot (.). Zinc coordination cysteine and histidine residues are coloured in pink and orange respectively. Hydrophobic residues are coloured green, polar/charged residues in blue, aromatic residues in purple, glycine/small residues in red and proline residues in yellow. **X** represents the acidic/hydrophobic residues important for the site-specific recognition of H3K4me0. MOZ amino acid residue numbers are denoted in red. Multiple alignments were conducted using the *MultAlin*, ExPasy based program.



[illegible]

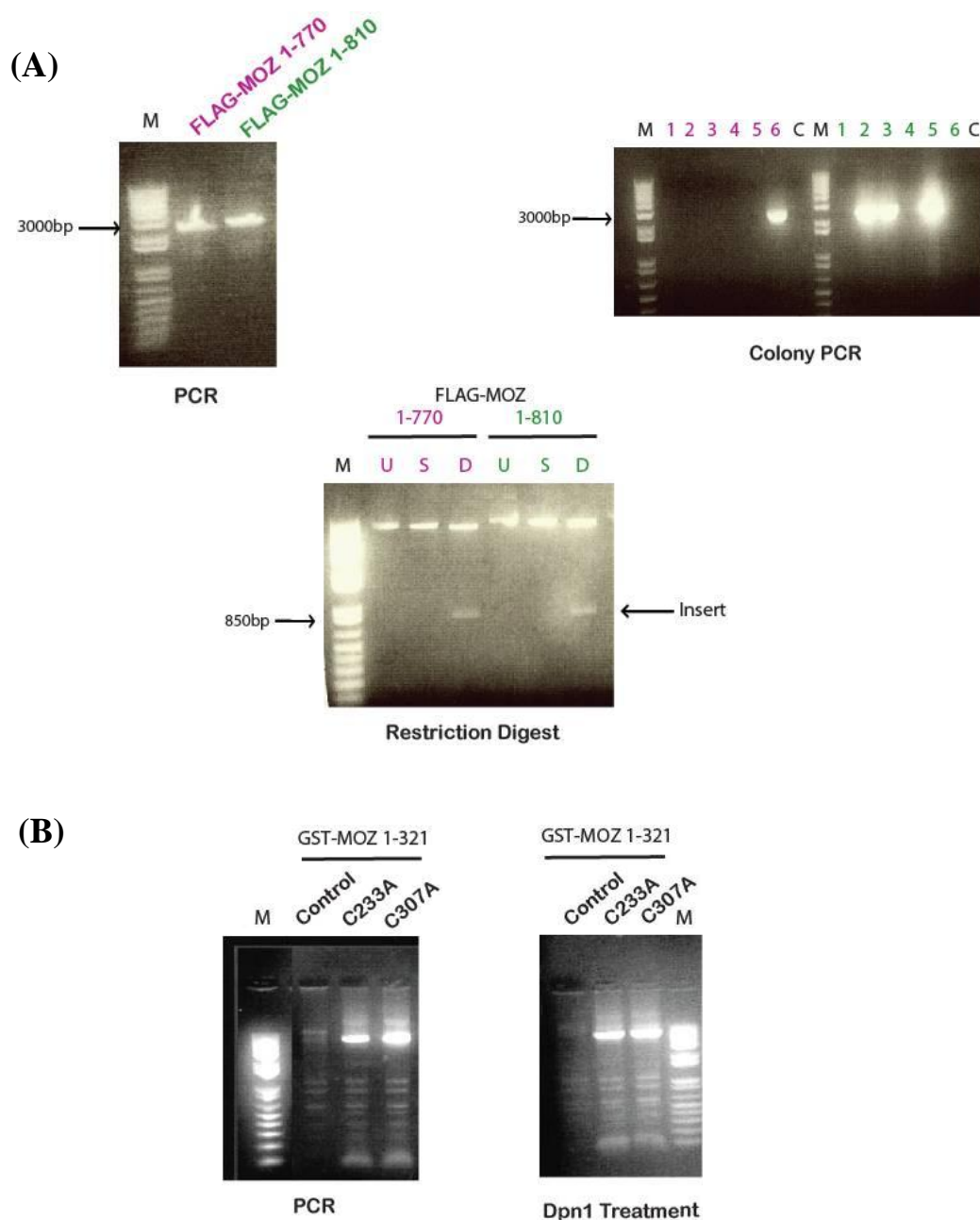
**Figure 3.3. Sequence alignment with the DPF family.** Alignments of individual PHD fingers from tandem motifs in the MYST and DPF families. Top and bottom alignments correspond to the first and second PHD fingers of the tandem motif respectively. Identical residues are marked with an asterix (\*), strongly similar residues denoted with a colon (:) and weakly similar residues denoted with a dot (.). Zinc coordination cysteine and histidine residues are coloured in pink and orange respectively. Hydrophobic residues are coloured green, polar/charged residues in blue, aromatic residues in purple, glycine/small residues in red and proline residues in yellow. **X** Represents amino acids implicated in DPF3b H4Kac binding that are not conserved in the MOZ PHD sequence. Multiple alignments were conducted using the *MultAlin*, ExPASy based program.



### 3.3. The N-terminal domains of MOZ bind to histone H3

To study the biochemical function of MOZ, a series of recombinant MOZ constructs were generated including a FLAG- or GST- tag. Oligonucleotides were designed according to the boundary requisite and used in PCR reactions to generate the DNA fragments needed for the construction of recombinant plasmids (Figure 3.4(A) - left panel). FLAG-tag and kozak sequences were also incorporated into oligonucleotide sequences used for cloning inserts into mammalian expression vectors. Next, a bacterial expression vector (pGEX) containing the sequence encoding a GST tag or a mammalian expression vector (pcDNA3.1+) were digested within their multiple cloning site with restriction endonucleases. Digested DNA fragments and vectors with compatible ends were then annealed together in a ligation reaction with T4 DNA ligase as described in the methods. Ligated DNA was transformed into DH5 $\alpha$  *E.coli*, plasmid preparations made and verified for the correct integration of the gene fragment by colony PCR (Figure 3.4(A) - right panel) and restriction digest (Figure 3.4(A) - bottom panel).

Site directed mutagenesis by PCR was used to create mutations in template DNA. This generated a full-length PCR product containing the mutation (Figure 3.4(B) – left panel). Methylated, non-mutated parental DNA was then digested with the Dpn1 restriction enzyme (Figure 3.4(B) – right panel, compare control lane to 2&3), DNA transformed into DH5 $\alpha$  *E.coli*, plasmid preparations made and verified for the correct mutation of the gene fragment by sequencing.

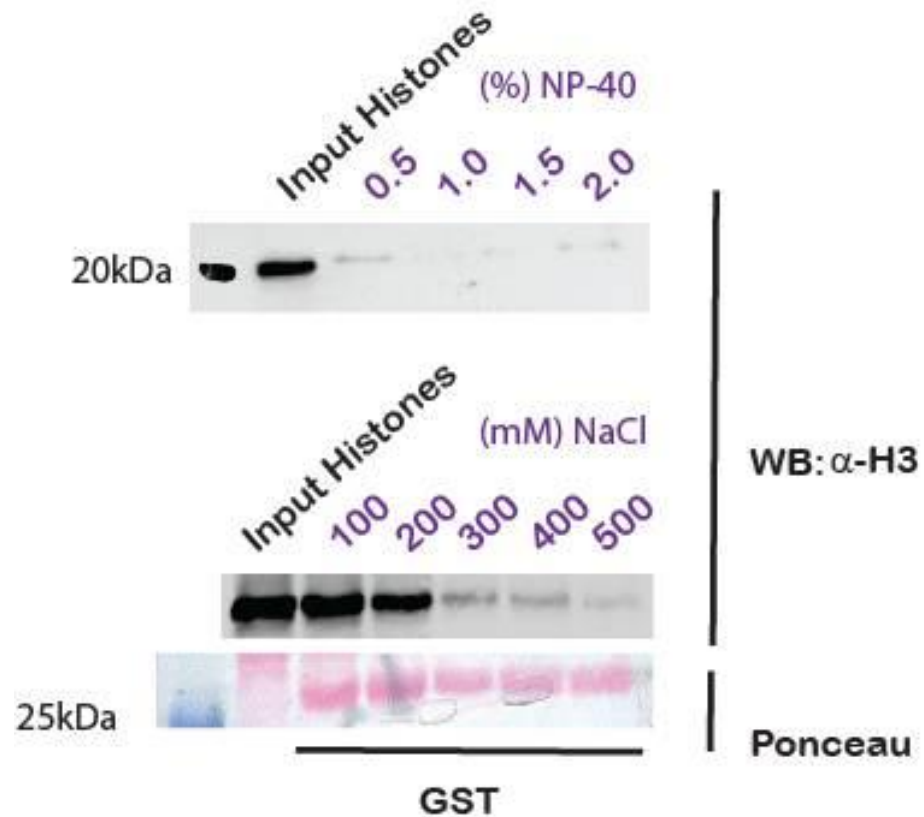


**Figure 3.4. Creation of GST-fusion and FLAG-constructs used in this study.**

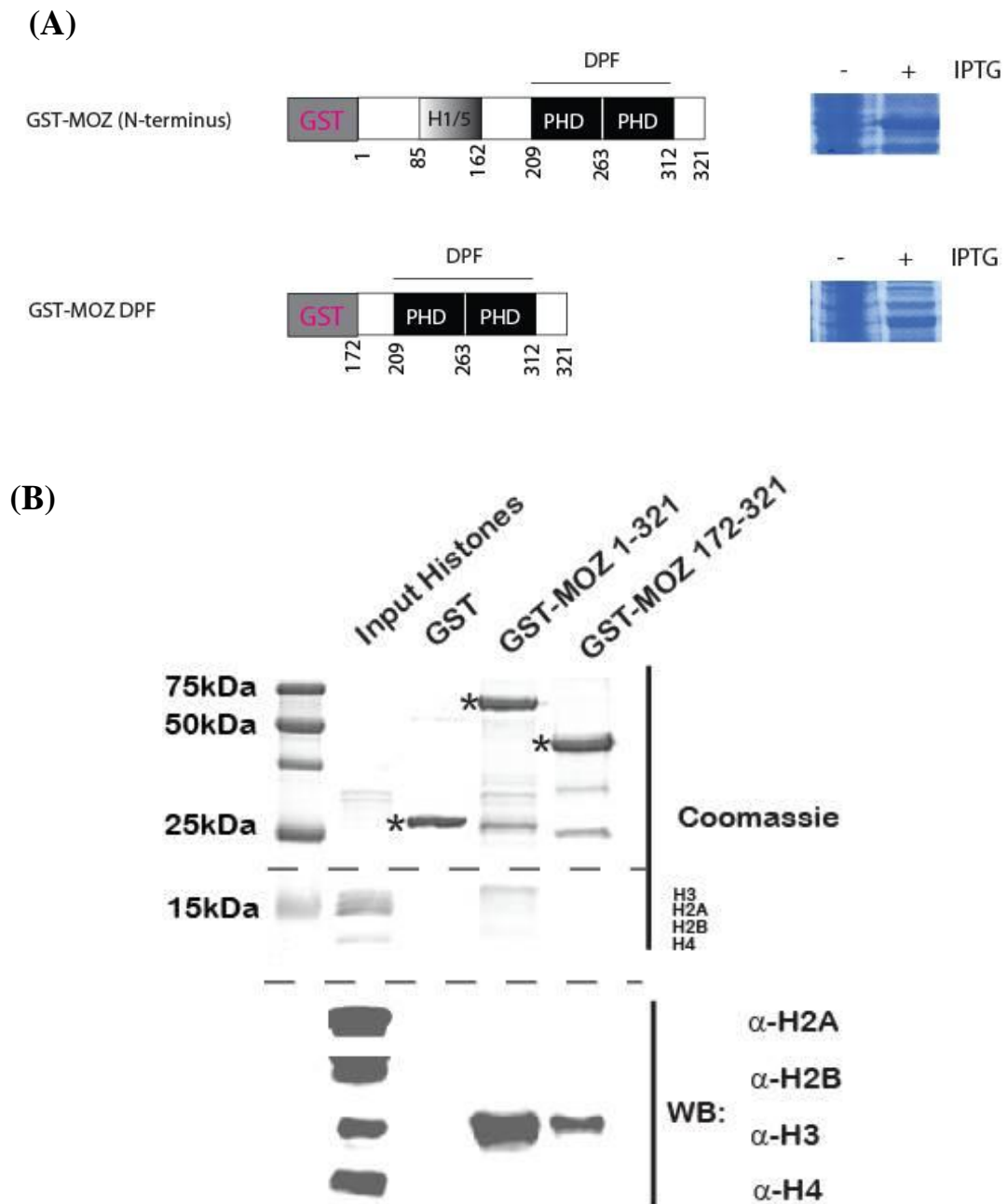
(A) Example of the standard cloning protocol, whereby desired products were amplified by PCR (top left panel) and inserted into a pre-digested vector. Newly synthesised clones were then screened for the correct insertion of product by colony PCR (bottom panel) and restriction digest (top right panel). Abbreviations; (M) marker, (C) control, (U) uncut, (S) single digest, (D) double digest. (B) Site-directed mutagenesis protocol was carried out as described in Materials and Methods.

A bacterial expression system was used in order to produce the proteins required for testing in GST-pulldown assays. Firstly, plasmid DNA encoding the GST-fusion protein was transformed into the Rosetta strain of *E. coli*. Rosetta host strains are BL21 derivatives designed to enhance the expression of eukaryotic proteins that contain codons rarely used in *E. coli*. These strains supply tRNAs for these specific codons on a compatible chloramphenicol resistant plasmid. Recombinant proteins were expressed by the addition of the IPTG inducer (Figure 3.6(A)) and subsequently, the host Rosetta cells were lysed to release the soluble protein. GST-fusion proteins were then purified via the high affinity of GST for glutathione using glutathione cross-linked sepharose beads. Protein levels were normalised to ensure equal loading in subsequent experiments (Figure 3.6(B)- top panel).

Our previous studies using MOZ fusion proteins had shown that MOZ can associate with chromatin and alter the acetylation of histones (Collins et al., 2006; Kindle et al., 2005). We thus assumed that the DPF might have a role in chromatin targeting via histone interaction. To assess that MOZ can interact with histones an *in vitro* GST pulldown assay was performed using the purified N-terminus GST-MOZ 1-321 or the DPF GST-MOZ 172-321 and core histones extracted from calf thymus. As a control, core histones were incubated with GST. Glutathione immobilised GST-MOZ fusion proteins were incubated with core histones, samples washed stringently to remove non-specific binding proteins and tested for pulldown (interaction) of histones by SDS-PAGE or immunoblotting.



**Figure 3.5. Establishing optimal conditions for the GST pulldown assay.** Pulldown assays followed by western blotting and immunodetection of histone H3 using the GST protein and calf thymus histone extracts. Buffer salt and detergent compositions were altered until the background binding of the GST protein tag to the core histones was eliminated.



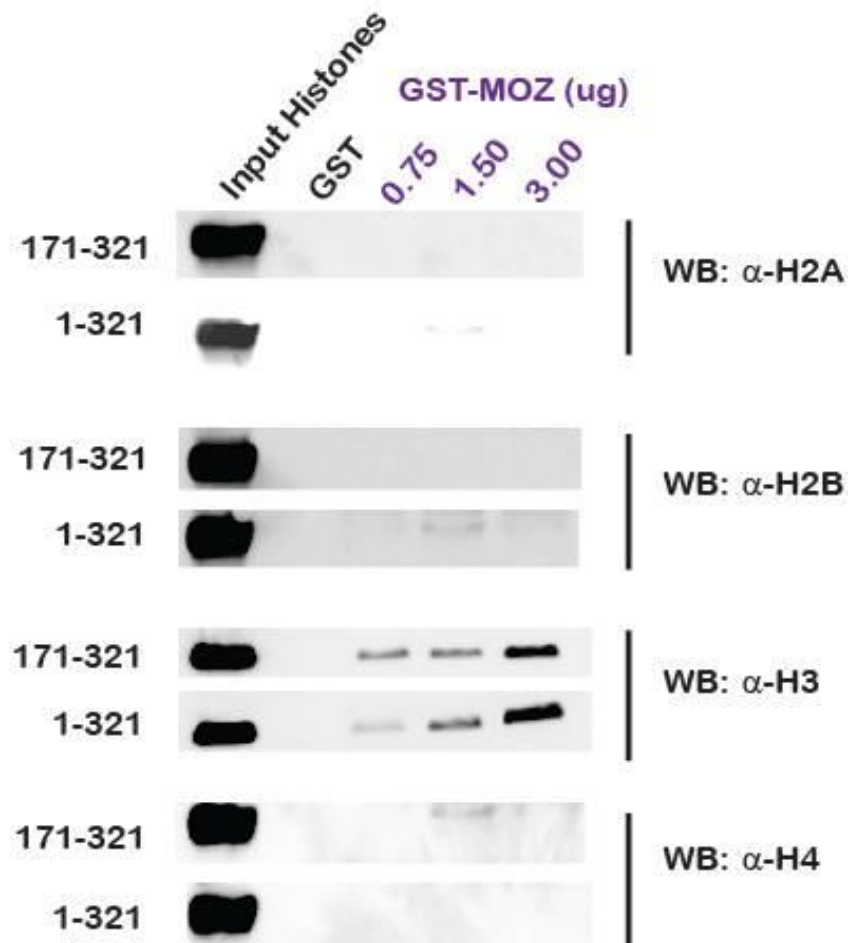
**Figure 3.6. The N-terminal domains of MOZ bind to Histone H3.** (A) Schematic representation of GST-MOZ fusion constructs used in subsequent assays and the expression of constructs in Rosetta *E. coli* bacteria following IPTG induction. (B) Pulldown assays followed by Coomassie staining or Western blotting and immunodetection of indicated histones using GST-MOZ fusion proteins and calf thymus histone extracts.

In preliminary experiments it was seen that the control GST protein non-specifically bound core histones when incubated in low salt buffers (Figure 3.5- lanes 2&3 middle panel). Ponceau staining indicates equal loading of GST protein (Figure 3.5- bottom panel). When salt was titrated into the buffer composition, binding of the GST protein to histones was lowered significantly (Figure 3.5- lanes 4-6 middle panel). Varying detergent concentrations in addition to this did not affect background binding further (Figure 3.5- lanes 3-6 top panel). Thus, optimal buffer conditions were established and set to high salt (500mM NaCl) and low detergent (0.5% NP-40) concentrations for all subsequent assays.

GST protein concentrations were normalised by SDS-PAGE and Coomassie staining to ensure equal amounts of protein was used throughout the assay (Figure 3.6(B)- top panel). Pulldown assays resulted in no interaction between the control GST protein and core histones, showing the interaction does not occur through the GST tag fused onto MOZ (Figure 3.6(B)-lane 3 middle panel). When histones and GST-MOZ 1-321 were incubated together, histone H3 was retained on the GST-conjugated beads suggesting that these two proteins interact directly *in vitro* (Figure 3.6(B)-lane 4 middle panel). In contrast, using Coomassie staining as a protein detection method did not identify an interaction between the MOZ-DPF (172-321) and histones (Figure 3.6(B)-lane 5 middle panel), thus the increased sensitivity of immunodetection was used to test this result instead. Consequently, pulldown assays were followed by the immunodetection of histones H2A, H2B, H3 and H4 (Figure 3.6(B)-bottom panel). Both MOZ 1-321 and MOZ DPF bind strongly to histone H3 in these experiments, however we noted that MOZ 1-321 showed a more robust binding. This suggests that the MOZ 1-171 sequence might stabilise the binding of or

provide additional contacts to histone H3. Interestingly, this sequence contains a region showing weak homology with histone H1 and H5. Under these conditions we were unable to detect binding with histones H2A, H2B and H4. This may be explained by the presence of a mixed population of PTMs within the core histone extract and will be analysed in more detail in subsequent assays.

The interaction of both MOZ constructs with histone H3 was confirmed by titration of the GST-MOZ protein into pulldown assays. This led to dose-dependent binding of histone H3 for both the N-terminus and the DPF domain alone (Figure 3.7). A weak interaction with histone H2B and H4 was also observed in the titration assays (Figure 3.7). Moreover, the binding to H4 was confirmed in subsequent experiments (See Figure 3.11)



**Figure 3.7. Dose-dependent binding of MOZ N-terminal domains to Histone H3.** Pulldown assays followed by Western blotting and immunodetection of indicated histones, using a titration of GST-MOZ fusion concentrations (0.75, 1.5 and 3  $\mu$ g) and 10  $\mu$ g calf thymus histone extracts.



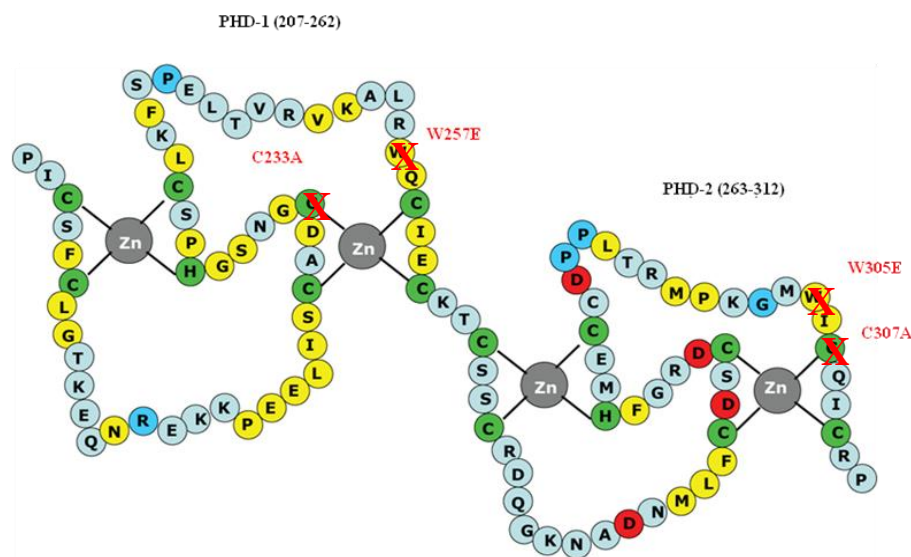
### **3.4. Binding of MOZ to histone H3 requires the integrity of the DPF**

To investigate whether the integrity of the MOZ DPF is necessary for histone binding, point mutations were introduced into the DPF sequence by site-directed mutagenesis. Residues selected for mutation are depicted on the schematic cross-brace structure, speculated to be adopted by a double PHD finger (Figure 3.8(A)). Firstly, two zinc-coordinating cysteine residues were mutated, one in each PHD finger (C233A & C307A). The cysteine residues do not interact with H3 directly but are required for the structural integrity of the PHD finger. Secondly, two conserved aromatic residues implicated in histone recognition and interactions in other PHD fingers, were mutated (W257E & W305E). All mutants were generated in the GST-MOZ 1-321 construct, expressed in Rosetta cells and purified using glutathione cross-linked sepharose beads as described in section 3.3. Protein concentrations were normalised by SDS-PAGE and Coomassie staining to ensure equal amounts of each protein were used in the assay. From this gel it is evident that slightly less GST-MOZ protein harbouring the W257E mutation was utilised in the subsequent assays, however this will be taken into consideration when interpreting latter results (Figure 3.8(B)). Using optimal buffer conditions, each of the PHD mutants, WT and the GST control was incubated with core histones to investigate whether an intact PHD finger is required for an interaction with histones. Pulldowns were subsequently subjected to immunoblotting with the primary antibody raised against H3 (Figure 3.8(C)). No interaction is observed between the GST control and histones (Figure 3.8(C)-lanes 2 & 8), but binding is evident between histone H3 and WT GST-MOZ 1-321,

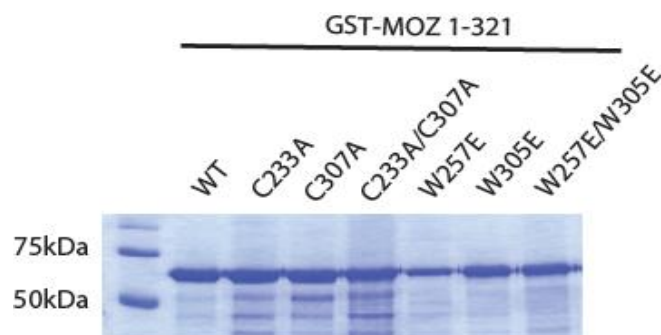
validating results observed in section 3.3 (Figure 3.8(C)-lanes 3 & 9 top panel). Ponceau staining indicated equal loading of proteins in each assay reaction (Figure 3.8(B) bottom panel).

Pulldown assay results with both the single cysteine and aromatic PHD finger mutants show a reduced interaction with histone H3 compared to wild-type (Figure 3.8(B)-compare lanes 3 & 9 with 4-5 & 10-11). It appears that the single aromatic mutant in MOZ PHD1 (W257E) has a greater negative effect on binding to H3 than the MOZ PHD2 (W307E) aromatic mutant (Figure 3.8(C)- compare lanes 10 & 11). However, as stated previously, slightly less of the MOZ W257E protein was used in the assay (Figure 3.8(B) and (C)- bottom panel) making it difficult to draw conclusions from this result. Further, when both PHD fingers contain point mutations the impact on H3 binding is reduced to a greater extent compared to proteins harbouring a single PHD mutation (Figure 3.8(C)- compare lanes 4-5 with 6 and 10-11 with 12). Thus, mutagenesis studies indicate that the binding of histone H3 by MOZ is mediated by the DPF domain and that the integrity of both PHD fingers is required for the interaction.

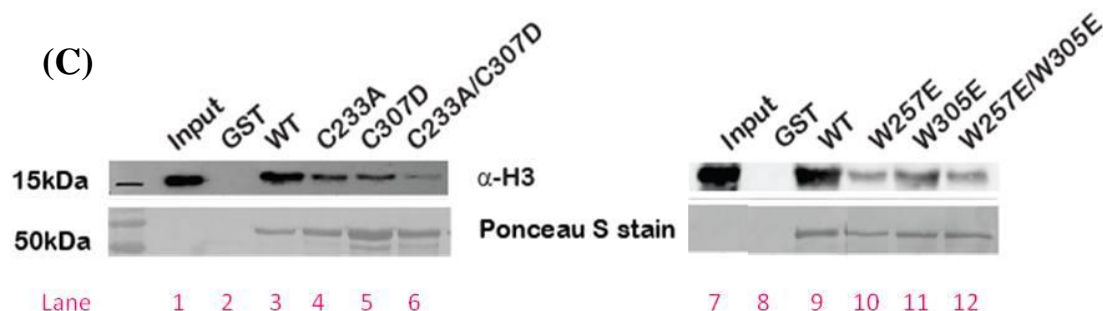
(A)



(B)



(C)

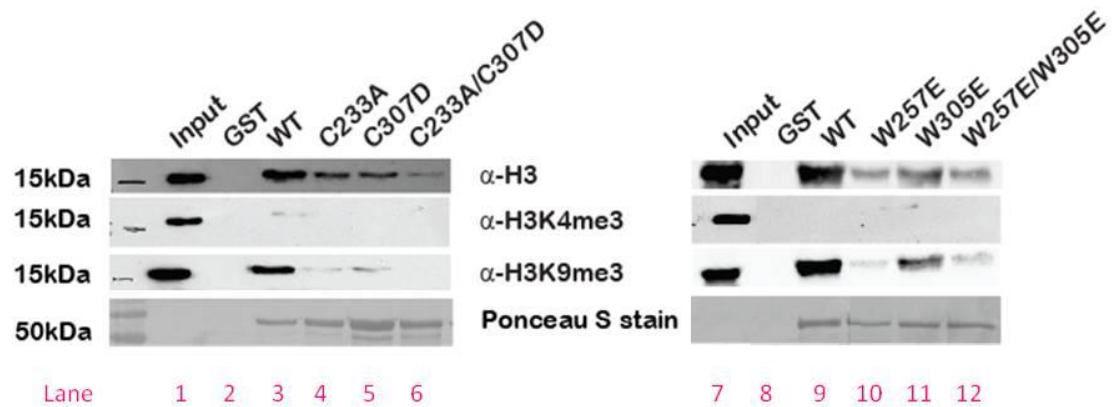


**Figure 3.8. Binding of MOZ to histone H3 requires the integrity of the DPF.**

(A) Schematic representation of the MOZ DPF cross-brace structure. Aromatic and zinc coordinating cysteine residues subjected to site-directed mutagenesis in GST-MOZ 1-321 are highlighted. (B) Normalisation of GST-MOZ 1-321 protein concentrations by SDS-PAGE and Coomassie staining. (C) Pulldown assays using GST-MOZ mutants and calf thymus histone extracts were followed by western blotting and immunodetection of histone H3.

### 3.5. Histone H3 and H4 PTMs alter the interaction with MOZ DPF

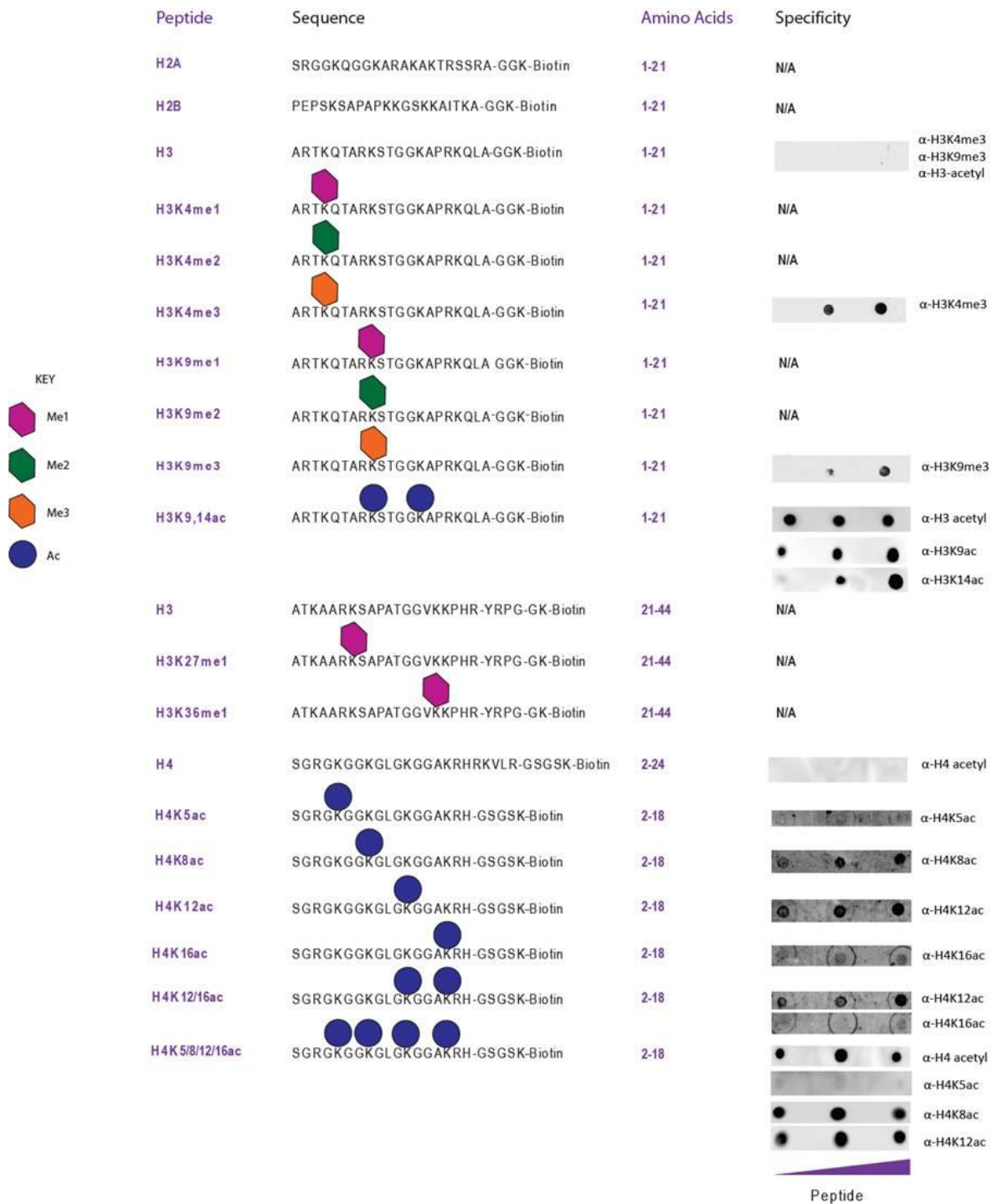
Having established that the MOZ DPF binds strongly to histone H3 and possibly to a lesser extent H4, we assessed how the histone modification status impacts on these interactions. As the H3K4 and H3K9 methylation status are major targets of other PHD domains, we reproduced the pulldown assays with the mutant GST-MOZ fusion proteins described in the previous sections, however this time blotting for the H3K4me3 and H3K9me3 modifications (Figure 3.9). Ponceau staining indicated equal loading of proteins in each assay reaction (Figure 3.9- bottom panel). As shown in Figure 3.9, WT MOZ 1-321 does not interact with the H3K4me3 modification; a PTM strongly correlated with an active transcription state (Figure 3.9- second panel). In contrast, WT MOZ shows a strong interaction with H3K9me3; a PTM strongly associated with heterochromatic regions (Figure 3.9- third panel). This suggests a possible role of the DPF in targeting MOZ to repressive chromatin or releasing MOZ from sites of transcriptional activation. As with histone H3 binding, single cysteine and aromatic PHD mutations reduce binding of GST-MOZ 1-321 to the H3K9me3 modification (Figure 3.9-third panel, compare lane 3 with 4-5 and 9 with 10-11). Further, to a greater extent, double PHD mutations completely abolish the interaction of MOZ 1-321 with histone H3 trimethylated at K9 (Figure 3.9- third panel, compare lane 3 & 9 with 6 & 12). These results indicate that the integrity of the MOZ DPF is required for histone PTM interaction and suggests that DPF binding is sensitive to the modification status of histone H3.



**Figure 3.9. Binding of MOZ to Histone H3 PTMs requires the integrity of the DPF.** Pulldown assays using GST-MOZ 1-321 mutants and calf thymus histone extracts were followed by western blotting and immunodetection of the indicated histone PTMs

Core histones are purified from cell extracts and thus, contain a mixed population of PTMs. Therefore, to determine the specificity of MOZ DPF binding further, biotinylated histone peptides modified at specific lysine residues were used in a pulldown assay with GST-MOZ fusions. Sequences of biotinylated peptides used in subsequent assays are listed in Figure 3.10 and where possible the presence of modification was verified using the dot blot technique. Increasing amounts of peptides were spotted onto nitrocellulose membrane and immunodetection was carried out with the indicated histone modification antibodies (Figure 3.10). Where tested, all modified peptides were validated by their corresponding antibodies and unmodified peptides were checked for the absence of modification by those antibodies raised against acetylated or methylated H3/H4 epitopes.

GST-MOZ 1-321 and DPF proteins were eluted from the glutathione-sepharose beads using the glutathione competitor and incubated overnight with the biotinylated histone peptides. Following this, biotinylated peptides were isolated via the high affinity of the biotin tag for Streptavidin using Streptavidin cross-linked sepharose beads. Samples were washed stringently as described in the materials and methods section and checked for the interacting GST-MOZ protein using SDS-PAGE and immunodetection with an antibody against the GST tag.



**Figure 3.10. Validation of the modification status of biotinylated histone peptides used in binding studies.** Dot blot analysis was carried out, where biotinylated peptides were spotted onto nitrocellulose membrane in increasing concentration increments (0.5, 1 and 2  $\mu$ g). Spotting was followed by immunodetection of the indicated histone post-translational modifications. N/A represents peptides not tested due to a lack of antibody against the specific modification.

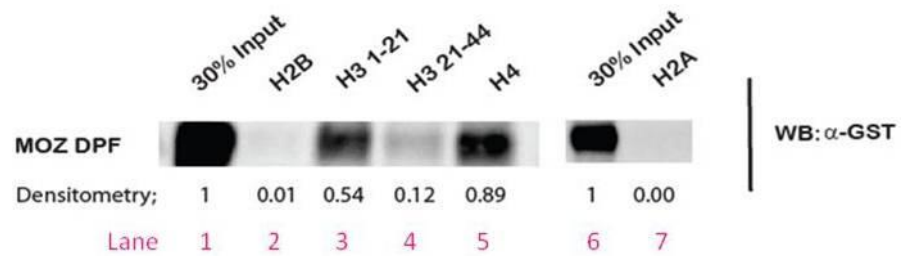
Initial experiments were performed to determine if the binding of MOZ DPF (194-323 - DPF boundaries were redefined) to histone H3 and H4 is mediated by the histone N-terminal tails. Firstly, to verify the results showing binding to core histones, this assay was carried out with peptides representing the N-terminus of each of the core histone tails and GST-MOZ DPF (Figure 3.11(A)). Results show that peptide H3 (1-21) interacts with the MOZ DPF, yet the internal H3 sequence (22-44) failed to bind (Figure 3.11(A) lanes 3 & 4). This suggests that the DPF binds the N-terminal 21 amino acids of the histone H3 tail. The H2A and H2B unmodified N-terminal tails also did not pulldown the MOZ DPF (Figure 3.11(A)-lanes 2 & 7), again reinforcing previous results. As shown more clearly in Figure 3.11, the MOZ DPF also strongly interacts with the N-terminal residues of histone H4 (2-24) (Figure 3.11(A)- lane 5). Subsequently, the pulldown assay was carried out with a series of H3 and H4 peptides acetylated and methylated at specific lysine residues (Figure 3.11(B(i))). Consistent with the absence of H3K4me3 binding in pulldown assays (Figure 3.9), both the MOZ 1-321 and DPF domains were unable to bind H3 1-21 containing the H3K4me3 modification (Figure 3.11(B(i)) compare lanes 3 & 5). Remarkably, dimethylation of H3K4 was bound equally as efficiently as the unmodified peptide (Figure 3.11(B(i)) compare lanes 3 & 4). This suggests that the DPF binding pocket can accommodate the size of dimethylated K4 but not with the added addition of a third methyl group. Alternatively, binding of H3K4me3 may be unfavourable as a consequence of steric repulsion, due to the increased cationic radius/charge of an additional methyl moiety.

In contrast to H3K4, trimethylation of H3K9 did not affect binding to either MOZ 1-321 or the DPF domain (Figure 3.11(B(i))- compare lanes 3 & 6). However,

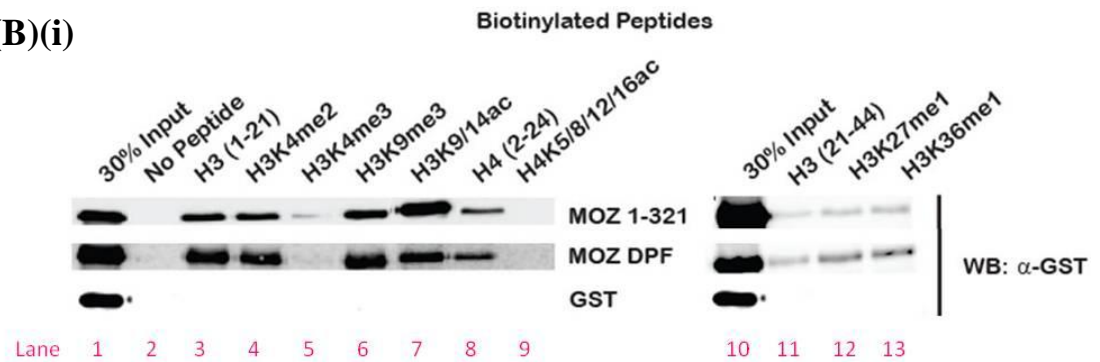


acetylation of K9 and K14 appeared to stimulate the interaction with MOZ 1-321 (Figure 3.11(B(i))- top panel, compare lanes 3 & 7), although this was not apparent with the DPF domain under similar conditions (Figure 3.11(B(i))- middle panel, compare lanes 3 & 7 (ii) compare lanes 2 with 4, 5 and 6). This suggests that acetylation of H3K9 or K14 may induce additional interactions with MOZ via the 1-171 sequence. Asymmetric dimethylation of H3R2 did not affect the binding of the DPF domain (Figure 3.11(B(ii))- compare lanes 2 & 3). Conversely, phosphorylation at H3S10 reduced the binding capacity of the MOZ DPF for the H3 N-terminal peptide (Figure 3.11(B(i))- compare lanes 2 & 7). This is possibly due to the appendage of a bulky, negatively charged phosphate group to the histone peptide, which can no longer be accommodated within the DPF binding cavity. Further, little or no interaction was observed with the internal histone H3 sequence (22-44) or the methyl PTMs in this region (Figure 3.11(B(i))-lanes11-13), consistent with MOZ DPF binding to N-terminal histone H3 (1-21) PTMs.

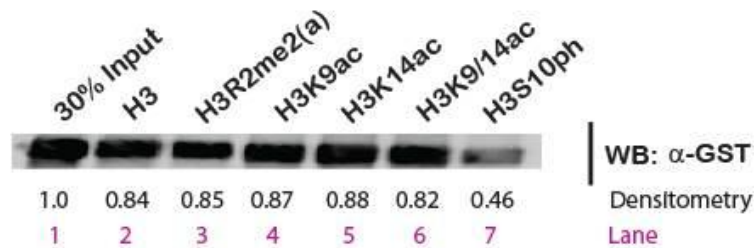
(A)



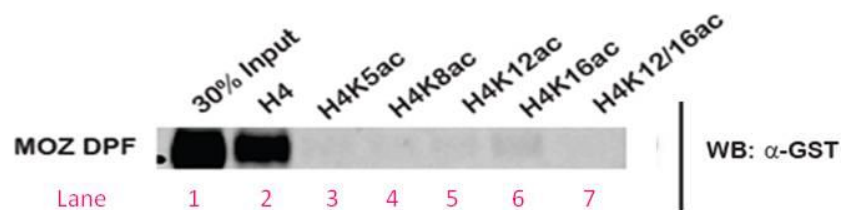
(B)(i)



(ii)

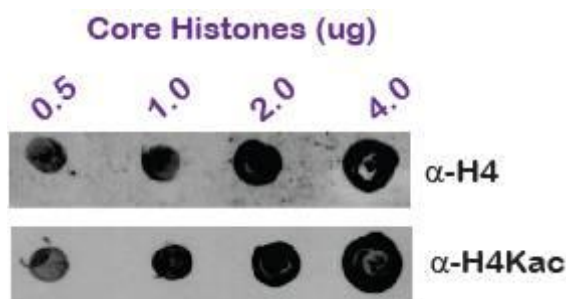


(C)



**Figure 3.11. Binding preference of the MOZ DPF: An intolerance of H3K4me3 and H4Ac.** Western blot analysis of histone peptide pulldowns with indicated GST-MOZ fusion proteins and biotinylated peptides (A) Pulldowns with core histone unmodified peptides. Densitometry readings were calculated, relative to the protein input, using the Image J software. (B)(i)(ii) Pulldowns with modified histone H3 and H4 peptides. (C) Pulldowns with histone H4 acetylated peptides.

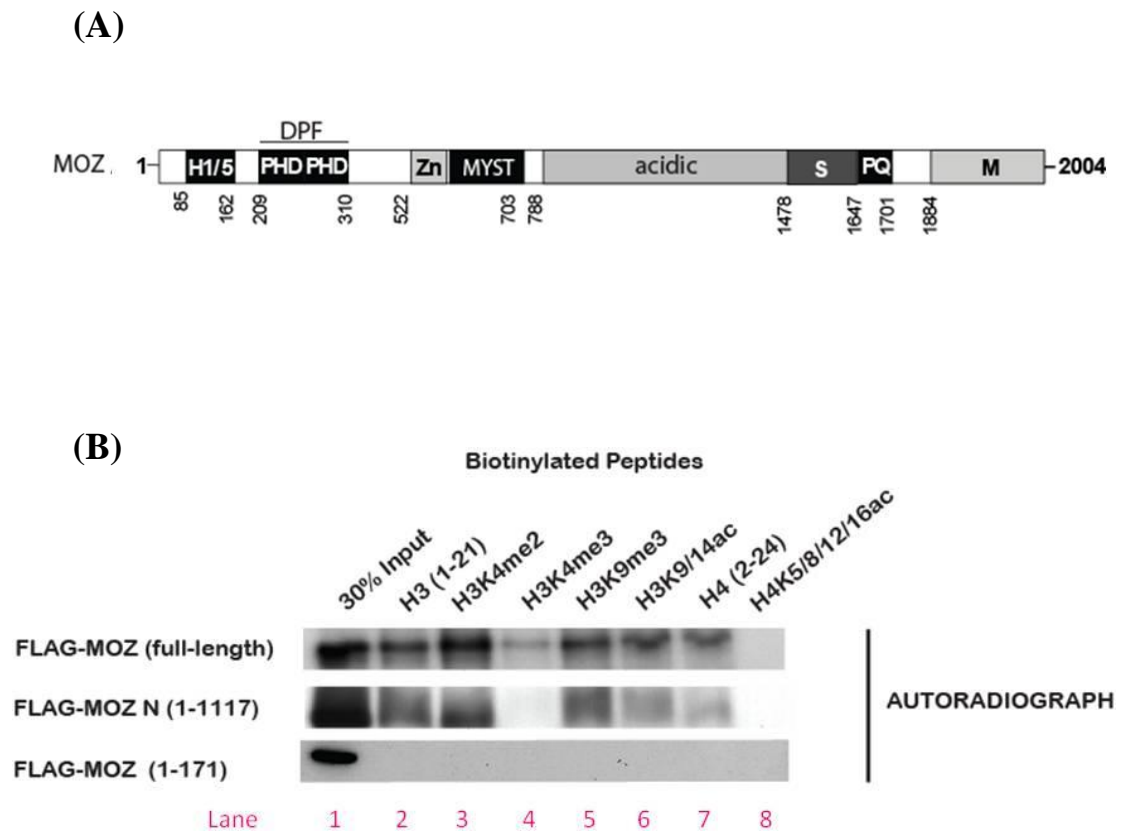
The 1-321 and DPF domains both showed strong binding to the histone H4 N-terminus, however both proteins were unable to interact with hyperacetylated histone H4 (Figure 3.11(B(i)) compare lanes 8&9) or further, with any H4 peptide mono/di-acetylated at each of the given target lysine residues (Figure 3.11(C) compare lanes 2 with 3-7). This indicates that H4 hyper-acetylation is incompatible with MOZ recruitment to chromatin. The binding of MOZ DPF constructs to histone H4, but not hyperacetylated H4, may provide an explanation for the rather poor binding to H4 observed with some preparations of core histones. Some commercial suppliers use sodium butyrate treatment (an inhibitor of HDAC activity) in their histone extraction protocol. This may mean that high levels of H4 are acetylated. To check for the presence of acetylated H4 in the histone preparations (Sigma), core histones were spotted onto nitrocellulose membrane and subjected to immunodetection with a pan-acetyl-H4 antibody. As shown in Figure 3.12, H4Kac was readily detected. This suggests that a high proportion of H4 may be acetylated and therefore unable to bind MOZ DPF. In contrast H3Kac may stimulate binding to MOZ-DPF.



**Figure 3.12. Core histone extracts contain acetylated histone H4.** Dot blot analysis. Core histones were spotted onto nitrocellulose membrane in increasing concentration increments (0.5, 1, 2 and 4  $\mu$ g). Spotting was followed by the immunodetection of histone H4 and pan-acetylation of histone H4.

MOZ is a modular protein consisting of several domains that could regulate or influence the binding specificity of the DPF domain (Figure 3.13(A)). Thus, to assess if the constructs used in our binding assays reflect the histone binding properties of the full-length protein, the full-length and breakpoint MOZ proteins were used in pulldown assays with the biotinylated peptides to investigate whether the binding patterns were similar. FLAG-MOZ and FLAG-MOZ N (which is truncated at the site of the oncogenic fusion protein breakpoint) were *in vitro* translated using a <sup>35</sup>S-methionine label, incubated with biotinylated peptides, immobilised onto Streptavidin beads and results visualised by SDS-PAGE and autoradiography. <sup>35</sup>S-methionine labelled proteins were used in this case as the larger MOZ proteins do not express and purify efficiently as GST fusion proteins. Input loading shows efficient *in vitro* translation (Figure 3.13(B)-lane 1). Full-length MOZ and MOZ N were found to associate with both histones H3 and H4 and displayed similar histone binding properties as the DPF and N-terminus of MOZ. Briefly, this involved intolerance for the activatory H3K4me3 mark and hyperacetylated H4 (Figure 3.13(B)-compare lanes 2 & 7 with 4 & 8).

Another protein expressed poorly as a GST-fusion was the GST-MOZ 1-171 (H1/5) protein. This construct was also *in vitro* translated with a <sup>35</sup>S-methionine label and used in peptide binding assays as described above. Input loading shows efficient *in vitro* translation of the protein (Figure 3.13(B)-lane 1). Interestingly, autoradiography shows no interaction of this module with any of the biotinylated peptides (Figure 3.13(B)-lanes 3-8). This further implements the MOZ DPF as the module responsible for histone recognition and suggests that the H1/5 domain may augment the DPF binding activity, rather than directly bind chromatin itself.



**Figure 3.13. Binding specificity of the full-length MOZ protein: An intolerance of H3K4me3 and H4Ac.** (A) Domain organisation of MOZ. (B) Autoradiograph analysis of histone peptide pulldowns with indicated *in vitro* translated <sup>35</sup>S-methionine labelled FLAG-MOZ, FLAG-MOZ N (oncogenic fusion protein breakpoint), FLAG-MOZ 1-171 and biotinylated histone peptides.

### 3.6. Both PHD1 and PHD2 are required to establish H3 PTM binding preference

The previous set of experiments established that MOZ can interact with H3 and H4 termini and that this is regulated by PTM status. To determine whether the two MOZ PHD fingers can function as two separate entities and carry out specific roles in governing the DPF binding pattern, GST-MOZ 194-263 (PHD1) and GST-MOZ 250-323 (PHD2) plasmid constructs were created (Figure 3.14(A)- left panel). These fusion proteins were expressed and purified as previously detailed. Both proteins were normalised to GST-MOZ DPF (194-323) using SDS-PAGE and Coomassie staining (Figure 3.14(A)- right panel) and subsequently used in binding assays with biotinylated histone peptides. Firstly, the binding specificity of the PHD fingers for the four histone peptides was investigated (Figure 3.14(B)). The MOZ DPF showed robust binding to histone H3 and H4 N termini, but not H2A, H2B or H3 (22-44) as in previous experiments (Figure 3.14 (B)- third panel). PHD2 showed a similar pattern binding unmodified H3 and H4 N-termini, but no binding to H2A or H2B (Figure 3.14 (B)- second panel). PHD1 was able to bind the H4 N-terminus but had a much reduced ability to bind the H3 N-terminus (Figure 3.14 (B)- first panel). This suggests that PHD1 contacts histone H4, whereas PHD2 can bind to both H3 and H4.

Next, PHD1 and PHD2 were incubated with a range of biotinylated peptides containing selected histone modifications and the observed binding patterns compared to those of the MOZ N-terminus and DPF (Figure 3.15(A)). Interestingly, again PHD1 alone showed no binding to H3 N-terminal peptides whether modified or

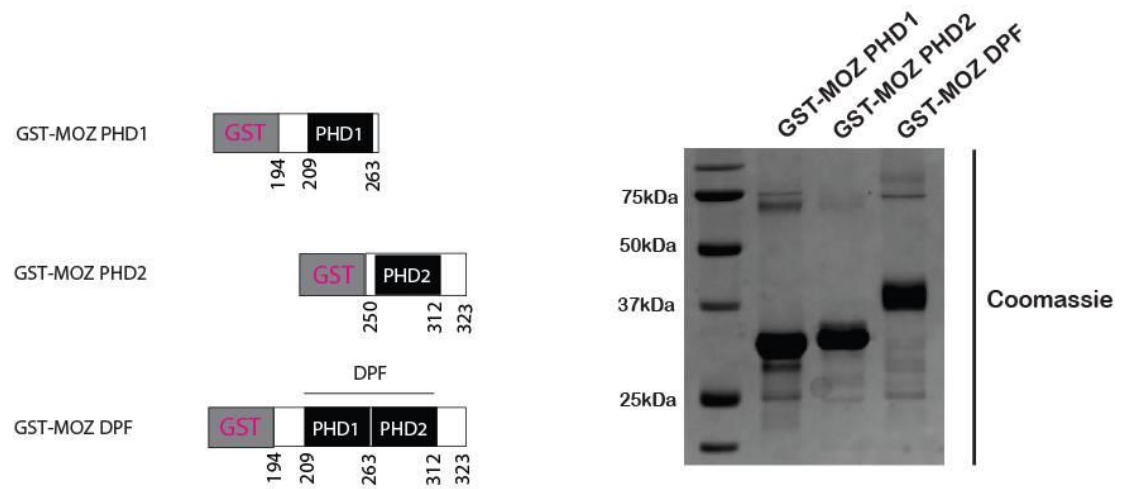
not (Figure 3.15(A)-lanes 3-7). A similar level of binding was seen to unmodified histone H4 as with the DPF, however hyperacetylation of H4 prevented this interaction (Figure 3.15(A)-compare lanes 8 & 9). To determine whether this effect is seen only with hyperacetylation rather than mono-acetylation of H4, we performed similar studies with mono-acetylated peptides. These results show that acetylation of K5, K8, K12 or K16 completely block the interaction of the DPF with the N-terminus of histone H4 under these conditions (Figure 3.15(B)-compare lanes 2 with 3-7). Surprisingly, similar results were observed using the H4 selective module PHD1, although in isolation this domain showed some ability to bind H4K16ac but not others. This may indicate that PHD1 and PHD2 can have functionality that is distinct from the composite DPF. This may be important in multicomponent complexes or after post-translational modification of MOZ within the DPF. In addition, no interaction was observed with unmodified histone H3 (22-44) or methyl PTMs in this region (Figure 3.15(A)-lanes 11-13).

PHD2 alone showed strong interactions with all N-terminal H3 peptides including the tri-methylated H3K4 mark (Figure 3.15(A)-lanes 3-7). Given the intolerance of the MOZ DPF protein for H3K4me3 binding, this result suggests that the PHD1 binding selection of H3K4me3 is dominant over that of PHD2 in the DPF composite. PHD2 also retained strong binding to histone H4, but like DPF and PHD1, H4 acetylation disrupted this interaction (Figure 3.15(A)-compare lanes 8 & 9). Again comparable to the DPF, a drastic reduction of binding was observed between PHD2 and any mono/di-acetylated H4 peptides (Figure 3.15(B)-compare lanes 2 with 3-7). This underscores complete intolerance for the binding of H4 acetylation PTMs. No interaction was observed with unmodified histone H3 (22-44) or methyl PTMs in

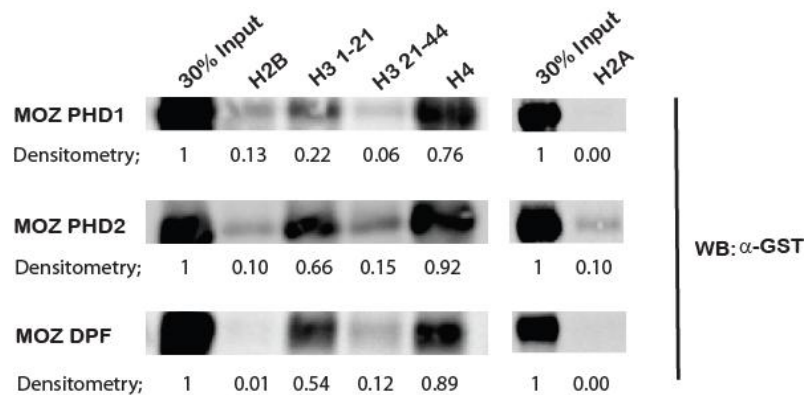
this region (Figure 3.18(A)-lanes 11-13), emphasising the specific binding of PHD2 and DPF to the very N-terminus (amino acids 1-21) of the histone H3 tail. These findings advocate that PHD1 and PHD2 functionally cooperate in an inter-dependent manner to determine the histone PTM binding selectivity of the DPF, in particular to the sensitivity to H3K4 trimethylation.



(A)



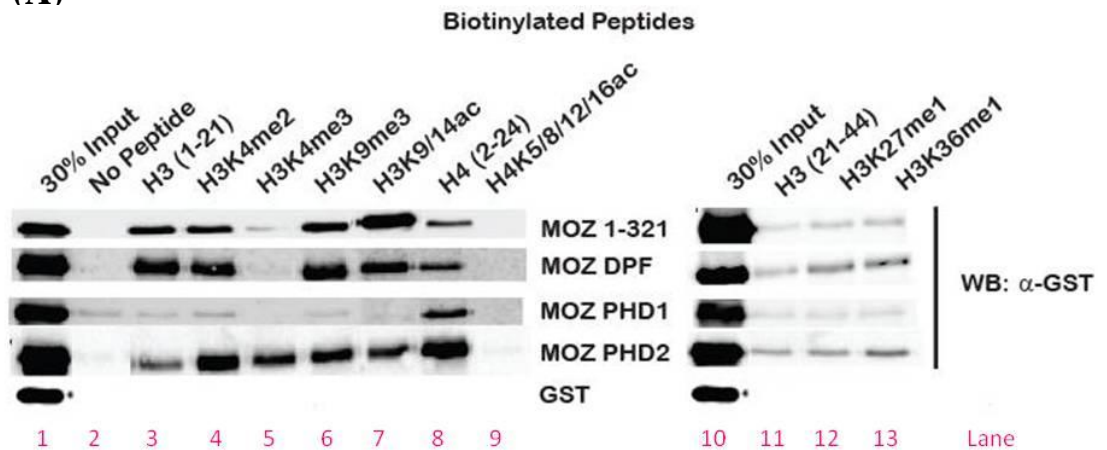
(B)



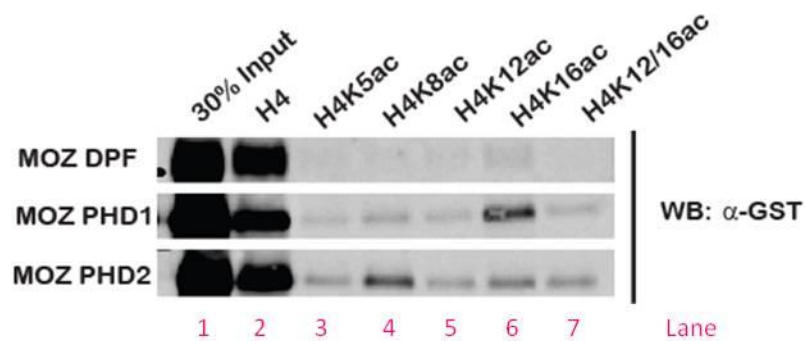
**Figure 3.14. Binding specificity of the individual PHD motifs of the MOZ DPF.**

(A) Schematic representation of the GST-MOZ split PHD finger fusions used in subsequent studies. SDS-PAGE followed by Coomassie staining was used to establish normalised input levels of each protein. (B) Western blot analysis of histone peptide pulldowns with GST-MOZ PHD split fusion proteins and core unmodified histone biotinylated peptides. Densitometry readings were calculated, relative to the protein input, using the Image J software.

(A)



(B)

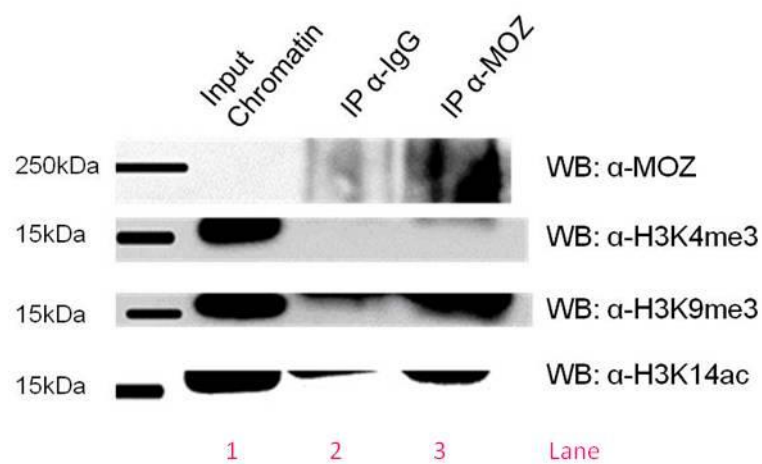


**Figure 3.15. Both PHD1 and PHD2 are required to establish H3 PTM binding preference.** Western blot analysis of histone peptide pulldowns with indicated GST-MOZ fusion proteins and biotinylated peptides. **(A)** Pulldowns with modified histone H3 and H4 modified peptides. **(B)** Pulldowns with histone H4 acetylated peptides.

### 3.7. MOZ is excluded from chromatin containing H3K4me3

The previous data indicated that *in vitro* the MOZ DPF can bind to both unmodified or acetylated histone H3 and unmodified histone H4, but is unable to interact with tri-methylated H3K4 and H4 that is mono- or hyperacetylated at the N-terminus. In order to investigate whether these *in vitro* findings are true for *in vivo* systems, ChIP assays and immunofluorescence studies were conducted.

HEK-293 cells have been shown to express a detectable level of endogenous MOZ (Kindle et al., 2005) and were consequently used in the *in vivo* studies. HEK-293 cells were cultured to confluency, the nuclear proteins and chromatin were cross-linked using formaldehyde and finally the chromatin was isolated using standard ChIP techniques (see materials and methods section 2.5.12). Endogenous MOZ was then immunoprecipitated from chromatin preparations and subsequently subjected to western blot analysis and immunodetection using histone-PTM specific antibodies. MOZ was indeed found to associate with chromatin, compared to IgG control, in these ChIP assays (Figure 3.16- panel 1, compare lanes 2&3). Western blotting with histone PTM-specific antibodies revealed that chromatin associated with MOZ contained acetylated H3K14 and H4K9me3 (Figure 3.16- panel 3&4, compare lanes 2&3), however MOZ did not coprecipitate with chromatin enriched in the H3K4me3 modification (Figure 3.16- panel 4, compare lanes 2&3). Input (10 %) samples were loaded (Figure 3.16- lane 1) to ensure the chromatin preparations were enriched with the specific histone PTMs being tested. Thus, it is likely that MOZ may be excluded from chromatin that is enriched in H3K4me3, but may be associated with chromatin containing H3K9/14 acetylation.



**Figure 3.16. MOZ is excluded from chromatin containing H3K4me3.** HEK-293 cells were crosslinked using formaldehyde and the chromatin isolated. Endogenous MOZ was immunoprecipitated from chromatin preparations and subsequently subjected to western blot analysis and immunodetection using histone PTM-specific antibodies.

### **3.8. MOZ colocalises with chromatin enriched in H3K9me3 and acetylated histone H3 PTMs**

The previous data indicates that MOZ is associated with chromatin containing H3Kac PTMs, but is excluded from H3K4me3 enriched chromatin. To investigate this further, we next examined the subcellular localisation of exogenous full-length FLAG-MOZ proteins and their association with endogenous histone PTMs using immunofluorescence microscopy.

To determine the subcellular localisation of MOZ, U2OS cells were seeded onto coverslips and transfected with the FLAG-MOZ expression plasmid. After 48 hours post-transfection cells were fixed, permeabilised and incubated with anti-FLAG antibody and a mouse Alexa594-conjugated secondary antibody. The nucleus was visualised with Hoechst 33258 stain which intercalates with DNA. As shown in Figure 3.17, FLAG-MOZ displays a punctate nuclear expression pattern, whereby the protein is concentrated into subnuclear foci and excluded from the nucleoli (Figure 3.17(A)). This localisation pattern is consistent with previous reports from our laboratory and is similar to that of the endogenous protein in cells that express MOZ (Kindle et al., 2005).

Next, FLAG-MOZ transfected U2OS cells were incubated with both anti-FLAG and anti-histone PTM antibodies and followed by incubation with a mixture of mouse Alexa594- and rabbit Alexa488-conjugated secondary antibodies. Both FLAG (red) and histone PTM (green) images were merged (yellow) to check for the nuclear localisation of the two markers. Cell nuclei were visualised with Hoechst 33258 stain (blue).

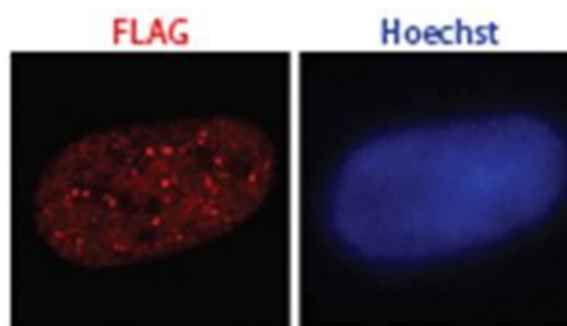
As shown in Figure 3.17, the FLAG-MOZ speckles do not overlap with H3K4me3 foci (Figure 3.17(B)- top panel), suggesting that MOZ is not localised to chromatic regions enriched with this modification. Indeed, the MOZ and H3K4me3 stains appeared to be largely mutually exclusive within the nucleus. This corroborates previous *in vitro* and *in vivo* experimental findings in this study. Although H3K4me3 is normally considered a marker of actively transcribed chromatin, it is not clear whether all transcribed genes contain this histone PTM. MOZ has been shown to activate certain genes, including *Hox* loci, so it does appear to have a function in gene activation. Our results may suggest that the function of MOZ in gene regulation either precedes or comes after the deposition of the H3K4 trimethylation.

Staining of U2OS cells for H3K9me3 revealed a mesh-like pattern in the nucleus with a number of larger foci enriched in this PTM (Figure 3.17(B)- bottom panel). These foci have been previously described as heterochromatic regions. In contrast to its relationship to H3K4me3, the FLAG-MOZ protein showed partial colocalisation with H3K9me3 enriched foci (Figure 3.17(B)- bottom panel). However, MOZ did not stain all H3K9me3 foci and in the typical example shown can be seen to be excluded from some of the larger foci. Thus, MOZ and H3K9me3 appear to be coincident in at least some regions of the genomes of interphase cells, consistent with our observation that this PTM does not displace MOZ from histone H3 *in vitro*.

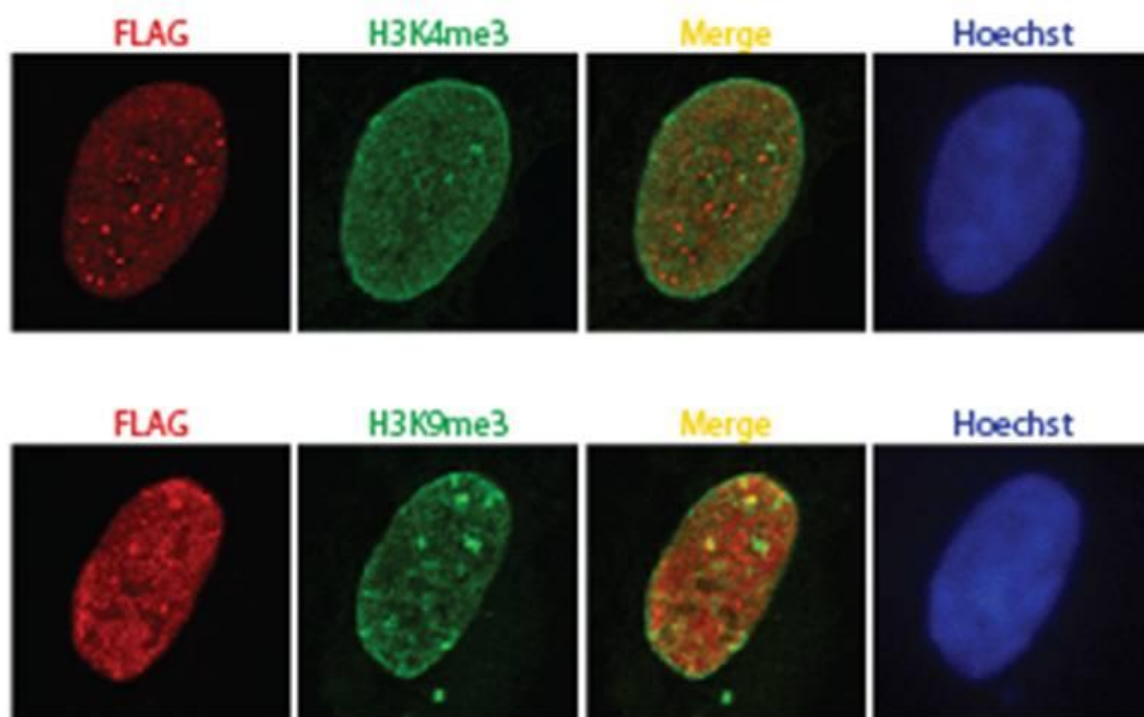
Previous assays showed that the MOZ DPF was able to bind and associate with H3Kac PTMs and as a result we next checked whether FLAG-MOZ colocalises with these marks using the immunofluorescence method. Anti-pan acetyl-H3 and anti-H3K9ac antibodies were successfully used in the subsequent studies. Various

dilutions of the H3K14ac antibody were tested, however signal above secondary control background was unattainable; thus the antibody was unsuitable for use in the immunofluorescence application. As shown in Figure 3.18, staining of U2OS cells for acetylated H3 revealed foci enriched with these PTMs. These foci showed excellent colocalisation with the FLAG-MOZ foci and indicate that MOZ is associated with regions enriched in acetylated H3 (Figure 3.18- top and middle panels). This is consistent with our *in vitro* data and also, with *in vivo* studies by Voss and Thomas showing that MOZ is specifically required for the H3K9 acetylation mark of transcriptionally active *Hox* gene loci (Voss et al., 2009). Similarly, FLAG-MOZ speckles were also found to show a good colocalisation with H3K9ac, although using this antibody fewer clearer foci were detected in U2OS nuclei (Figure 3.18- bottom panel). Thus, in summary it is likely that MOZ colocalises to and associates with chromatin enriched in H3 lysine acetylation marks.

(A)

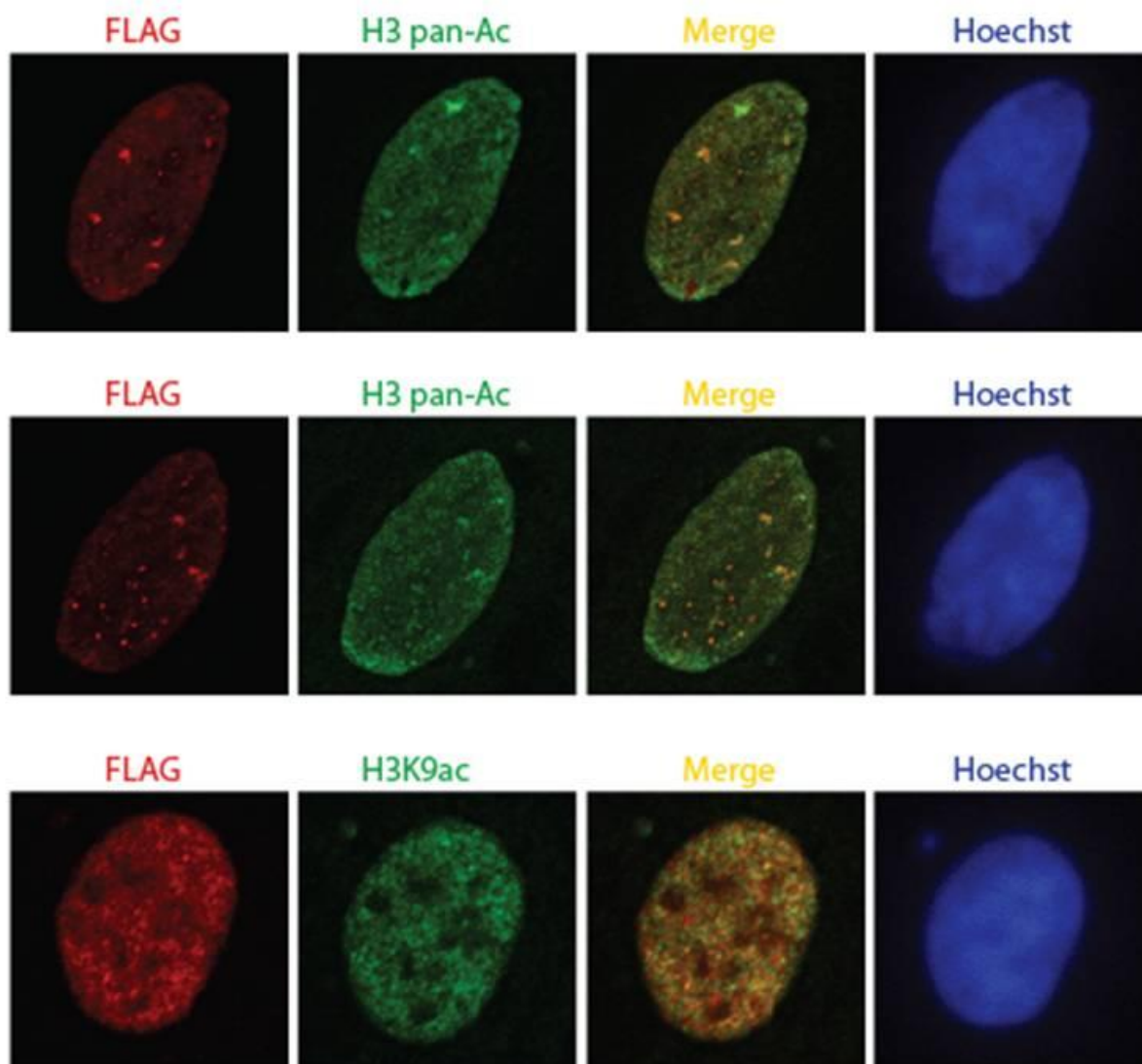


(B)



**Figure 3.17. Exclusion of MOZ from H3K4me3 enriched chromatin.** U2OS cells were seeded onto coverslips and transfected after 24 hours with expression plasmids encoding FLAG-MOZ. After 48 hours post-transfection, cells were fixed in 4 % paraformaldehyde, permeabilised and stained with antibodies raised against (A) mouse anti-FLAG (red) or (B) mouse anti-FLAG and the indicated rabbit anti-histone PTM (green). Staining was followed by incubation in (A) anti-mouse Alexa594 or (B) anti-mouse Alexa594 and anti-rabbit Alexa488 secondary antibodies. Merged scans (yellow) highlight colocalisation or exclusion and Hoechst 33258 (blue) was used as a stain of DNA.





**Figure 3.18. Colocalisation of MOZ with acetylated H3 enriched chromatin.** U2OS cells were seeded onto coverslips and transfected after 24 hours with expression plasmids encoding FLAG-MOZ. After 48 hours post-transfection, cells were fixed in 4 % paraformaldehyde, permeabilised and stained with antibodies raised against mouse anti-FLAG (red) and the indicated rabbit anti-histone PTM (green). Staining was followed by incubation in anti-mouse Alexa594 and anti-rabbit Alexa488 secondary antibodies. Merged scans (yellow) highlight colocalisation or exclusion and Hoechst 33258 (blue) was used as a stain of DNA.

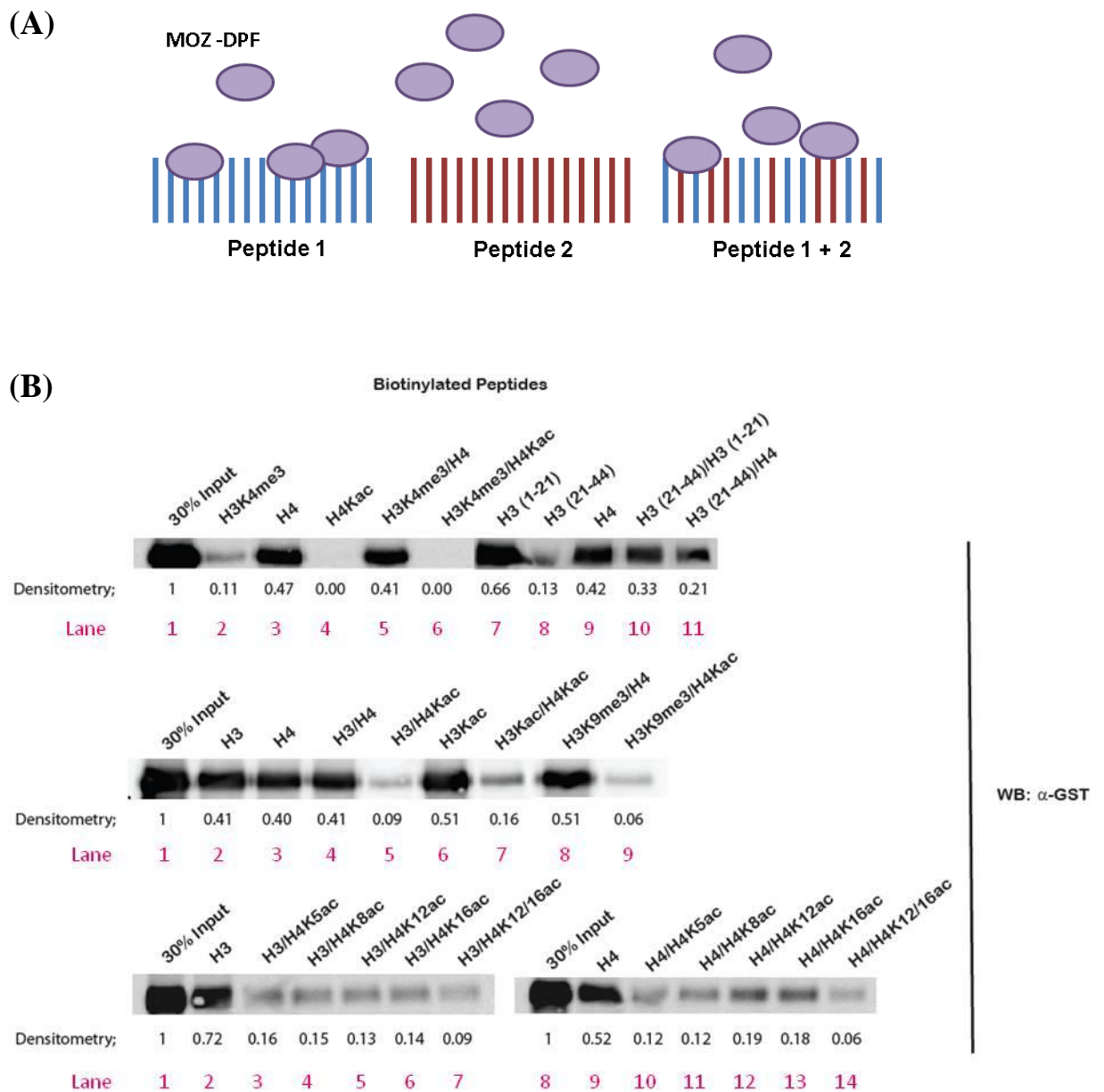
### 3.9. The MOZ DPF facilitates simultaneous reading of H3 and H4 PTMs

The coexistence of chromatin modifications on one or two tails of the same nucleosome suggests that it is likely that chromatin regulators can sense multiple PTMs cooperatively or simultaneously. To investigate how combinations of H3 and H4 PTMs influence MOZ DPF engagement, we devised an assay using equimolar mixtures of immobilised histone peptides. Schematic representation of the assay is depicted in Figure 3.19(A), where multiple peptides are immobilised onto Streptavidin-coupled Dynabeads®. Using single peptides, MOZ DPF showed similar binding patterns to H3/H4 PTMs as before, whereby strong binding was observed with H3 (1-21) (Figure 3.19(B) top panel- lane 7), H4 (top panel- lane 3 & 9), H3Kac (middle panel- lane 6) peptides and little or no binding was seen with H3 (21-44) (top panel- lane 8), H3K4me3 (top panel- lane 2) and H4Kac (top panel- lane 4).

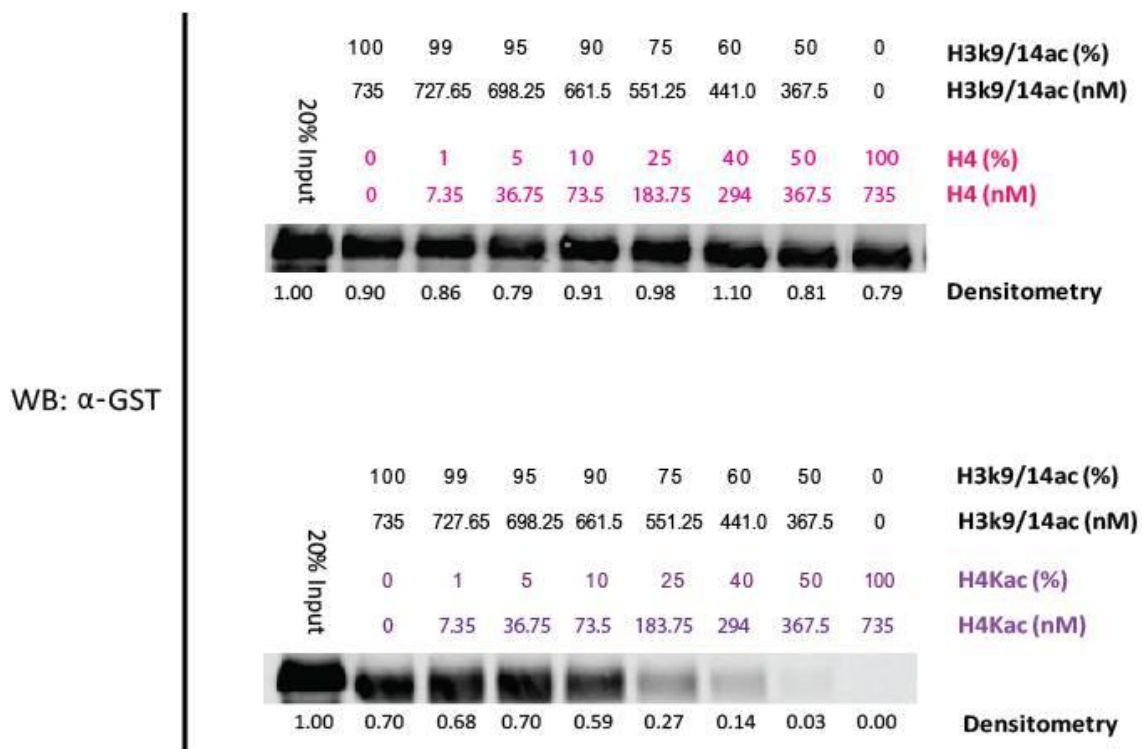
No increase in MOZ DPF binding was observed by providing a mixture of unmodified H3 and H4 peptides (Figure 3.19(B) middle panel- compare lanes 2&3 with 4), suggesting that the binding is not cooperative or allosteric. Notably, the availability of unmodified H3 or H4 peptide rescued the lack of MOZ DPF binding to histone H3 (22-44) (Figure 3.19(B) top panel- compare lanes 7-9 with 10 & 11) and H3K4me3 (top panel- compare lanes 2 & 3 with 5). This suggests that despite trimethylation of histone H3K4 alone excluding MOZ DPF from histone H3, the modification acts as a recessive exit signal, which does not affect the simultaneous binding to PTM residues on neighbouring histone tails or nucleosomes. In contrast, the addition of an acetylated H4 peptide into these assays appeared to have a

dominant inhibitory effect, as it was observed to absolutely prevent binding of the MOZ DPF to unmodified H3 (middle panel- compare lanes 2 & 5), H3Kac (middle panel- compare lanes 6 & 7) and H3K9me3 (middle panel- compare lanes 8 & 9). Furthermore, this dominant inhibitory effect was also seen with histone H4 peptides harbouring individual acetylation marks, where binding of the DPF to unmodified histone H3 and H4 was prevented in the presence of any mono-acetylated H4 peptide (bottom panel- compare lanes 2 with 3-7 and 9 with 10-14 respectively). These results suggest that DPF can simultaneously engage PTMs on different histones, at least in the context of peptide surrogates.

To further confirm the inhibitory effect of H4 acetylation upon binding to other recognised PTMs, peptide binding assays were carried out using H3K9/14ac and H4Kac peptides mixed in differing molar ratios with the GST-MOZ DPF protein. Results show a dose-dependent disruption of MOZ DPF binding to H3Kac by H4Kac, but not by H4 (Figure 3.20). Densitometry values emphasise a 16% reduction in DPF-H3Kac binding at a low 90:10 percentile ratio of H3Kac:H4Kac and a dramatic 62% decrease at a 75:25 ratio. This underscores the potency of the H4Kac inhibitory effect, where even at low concentrations of H4Kac and an excess of H3Kac, MOZ DPF binding to PTMs is diminished.



**Figure 3.19. The MOZ DPF facilitates simultaneous reading of H3 and H4 PTMs.** (A) Schematic representation of simultaneous peptide binding assays. (B) Peptide binding assays were conducted using either single biotinylated peptides or equimolar mixtures of two peptides and the GST-MOZ DPF protein. Western blot analysis and immunodetection using anti-GST followed. Densitometry values were calculated using the Image J software.



**Figure 3.20. Dose-dependent disruption of MOZ-DPF binding to H3Kac by H4Kac.** Peptide binding assays were conducted using H3K9/14 and H4 or H4 acetylated peptides mixed in differing molar ratios (a total of 735 nM peptide(s) was used). Peptides were incubated with 50 nM GST-MOZ DPF protein in 750  $\mu$ l binding buffer. Incubation with Dynabeads<sup>®</sup>, western blot analysis and immunodetection using anti-GST followed. Densitometry values were calculated, relative to the protein input, using the Image J software.

This advocates a possible model where MOZ complexes are likely to induce acetylation of H4, which consequently acts as a dominant exit signal from chromatin, even in the presence of H3 PTMs that stimulate MOZ recruitment. Alternatively, H4Kac could function as a signal preventing MOZ from docking to chromatin enriched in this modification in the first instance. It is difficult to envisage how an acetyl-H4 mark can block the interaction of MOZ DPF with other PTMs when it itself cannot bind. It is possible that an H4Kac group may ‘rest’ in the DPF binding site without forming covalent contacts with residues lining the pocket. This would suggest a transient and dynamic binding state, whereby the H4Kac group would temporarily block access to the reader module. Stringent washing steps in the *in vitro* assays may be rigorous enough to dislodge the H4Kac from the binding site; possible reasoning for a negative result for binding in these assays. Alternatively, every protein-ligand complex dissociates at a random time, causing a variable off-rate of the ligand dissociating from the protein. It is possible that the H4Kac peptide may bind briefly to the MOZ DPF protein and that complex dissociation exhibits rapid off-rate kinetics, causing the H4Kac peptide to swiftly dissociate from the DPF binding site. This could explain why little or no H4Kac modification was found in complex with the MOZ DPF in the peptide binding assays after stringent washing and further, how H4Kac was able to block the interaction of MOZ DPF with other PTMs. Dissociation binding experiments and structural crystallisation of the MOZ DPF in complex with an H4Kac peptide would aid the characterisation of the molecular mechanism by which H4Kac inhibits interaction with other PTMs.

### 3.10. The MORF DPF exhibits a similar binding pattern to MOZ DPF

MORF (MYST4) is another member of the MYST family of lysine acetyltransferases and is the most closely related member to MOZ in terms of its modular consistency and sequence homology. Alignment of the MOZ and MORF PHD fingers underscores the extent of this sequence similarity (Figure 3.21 (A)), where 87 % of amino acid residues are conserved (denoted with \*) and 95 % are strongly similar (denoted with :). As only 5 % of residues among the two DPF sequences are weakly similar or different (indicated with . or – respectively), it was expected that these two highly homologous domains would exhibit a similar specificity and binding function towards histone PTMs.

In order to test this hypothesis, the MORF DPF (amino acids 201-330) DNA fragment was cloned into the GST-expression plasmid, the protein expressed, purified and utilised in binding assays with modified biotinylated histone peptides (Figure 3.21 (B)). The binding pattern of MOZ DPF (Figure 3.11(B)- middle panel) in comparison to the MORF DPF (Figure 3.21(B)) is extremely similar. Like MOZ, MORF DPF interacts with both unmodified histone H3 and H4 peptides (Figure 3.21(B)- lanes 3 & 8). However, the MORF DPF shows a robust preference for unmodified histone H4 (Figure 3.21(B)- compare lane densities 3 & 8). When tested with histone H3/4 peptides acetylated and methylated at specific lysine residues, the two distinctive findings obtained for MOZ DPF binding were observed. Firstly, the MORF DPF showed diminished binding to H3K4me3 compared with a tolerance for H3K4me2 (Figure 3.21(B)- compare lanes 4 & 5), suggesting that the MORF DPF

binding pocket was also unable to sterically accommodate a third methyl group. In addition, the MORF DPF was unable to interact with hyperacetylated histone H4 (Figure 3.21(B)- compare lanes 8 & 9), implicating acetylation of H4 as an exit or prevention signal for MORF from chromatin. These results indicate that the homologous MOZ and MORF DPF domains exhibit very similar histone binding patterns, although MORF may display a particular preference for interaction with unmodified histone H4. This suggests that both enzymes are possibly targeted to or away from chromatin enriched in similar PTMs.





### 3.11. Quantitative analysis of MOZ DPF binding to unmodified or acetylated H3/H4 peptides

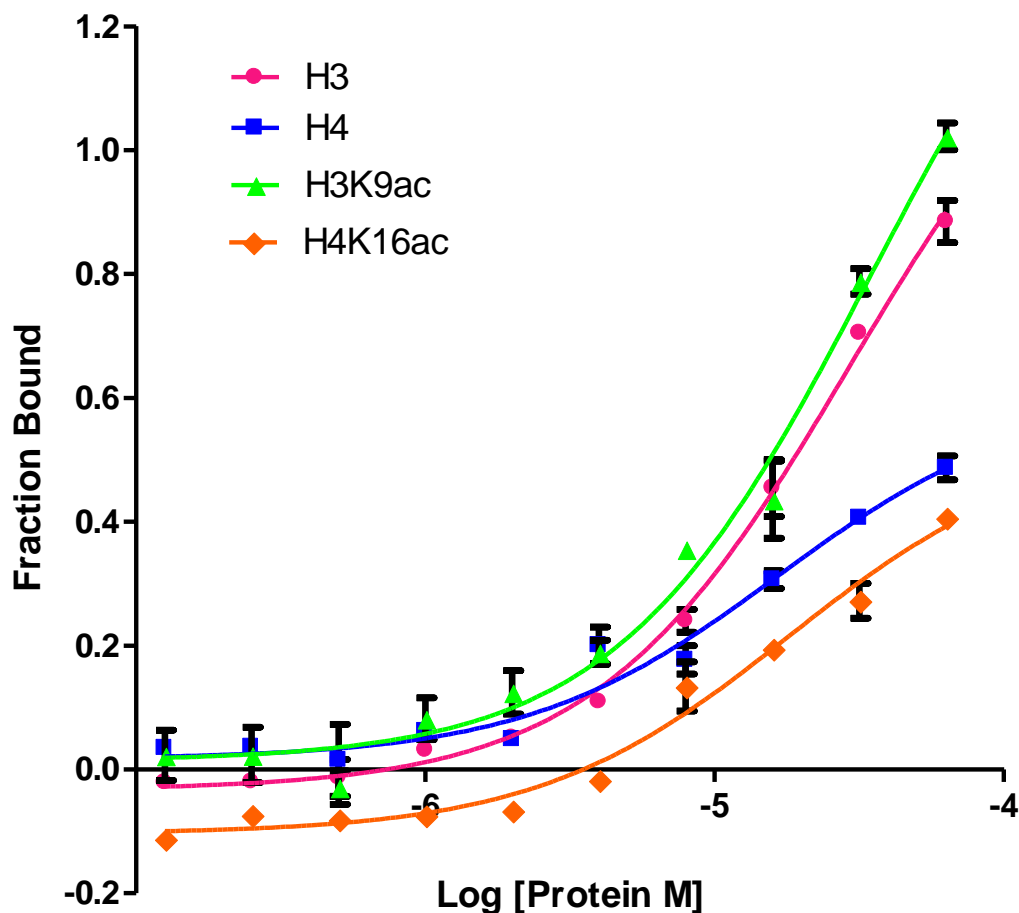
Biotinylated peptide binding assays are a useful tool to rapidly profile the interactions of MOZ DPF with histone PTMs. Densitometry readings yielded from such assays are useful in giving an indication of binding preferences. In order to measure binding preference for specific histone PTMs more quantitatively, fluorescence polarisation anisotropy was carried out as described in the Materials and Methods section. This method measures changes in the proportion of polarised light due to the reduced tumbling of a fluorescently labelled peptide upon interaction with a larger protein. Conditions for fluorescence polarisation were optimised to establish the background levels produced by different peptide concentrations. Optimal conditions were established by titration assays using different amounts of GST-MOZ DPF. Finally, assays were performed for GST-MOZ DPF binding to fluorescein-labelled H3 (1-21) and H4 (2-24) peptides. To obtain binding affinities, a constant concentration of each histone peptide substrate (1 nM) was incubated with increasing concentrations of GST-MOZ DPF (0.125 - 64  $\mu$ M). The fraction of peptide bound was then calculated relative to a background fluorescence reading.

As shown in Figure 3.22, polarised light due to reduced tumbling of the unmodified H3 and H4 peptides was increased upon the addition of increasing amounts of GST-MOZ DPF. The slope of the gradient for the histone H3 plot (pink) is steeper than that observed for the histone H4 peptide (blue). In addition, the overall fraction bound at high concentrations of MOZ-DPF is substantially higher for the H3 peptide compared to the H4 peptide (Figure 3.22), emphasising that the MOZ DPF

binds unmodified histone H3 with a greater affinity than histone H4. Quantitative binding measurements lend support to the apparent specificity of the MOZ DPF for the unmodified H3 peptide compared to the H4 peptide; a result not established by peptide binding assays. Unfortunately, binding affinities ( $K_d$ -dissociation constant) could not be calculated accurately using these results, as binding saturation was not reached at the maximal possible concentration of GST-MOZ DPF. An alternative method that might require less protein is isothermal titration calorimetry, which may be performed in future studies.

The binding studies with biotinylated peptides suggested that acetylation of H3 and H4 had different effects on the binding of MOZ DPF. The acetylation of H3K9 and H3K14 resulted in increased binding to GST-MOZ 1-321, whereas this was less apparent with GST-MOZ DPF alone. Conversely, monoacetylation of K5, K8, K12, K16 or hyperacetylation of all four residues within the H4 peptide completely eradicated the interaction with MOZ 1-321 or the DPF (Figure 3.11). To further investigate this binding preference shown by the MOZ DPF, polarisation experiments were set up using fluoroscein-labelled H3K9ac (1-21) and H4K16ac (2-24) peptides. As shown in Figure 3.22, binding of the DPF to H3K9ac (green) was increased in comparison to unmodified H3 (pink) (Figure 3.22). In contrast, acetylation of H4K16 (orange) reduced the level of binding observed in the polarisation assay, when compared to the H4 peptide (blue) (Figure 3.22). This result indicates that the MOZ DPF has a lower affinity for the H4K16ac peptide. Due to the high cost of fluoroscein-labelled peptides, we were unable to perform a control to establish whether this low level of increased polarised light is due to weak binding, or due to background. Background due to non-specific interactions could be established

by incubating the DPF with a negative control peptide, i.e a sequence for which the DPF has no affinity. However, on balance, it is likely that the acetylated H4 peptide has a weak affinity for DPF that is below the level of detection in the biotin-peptide pulldown assays. This could provide an explanation for the dominant inhibitory affect over the binding of other histone PTMs, whereby acetylated-H4 peptides block the ability of the DPF to bind H3. It will be necessary to generate a crystal structure for DPF/H4 complexes to establish whether H3 and H4 N-terminal tails dock at a similar or distinct site(s) on MOZ DPF. It is possible that even a weak affinity for acetylated-H4 tail might induce a conformational change to displace the H3 tail. Given the apparent difference in binding strengths, displacement of H3 by H4Kac through competitive binding seems a less likely explanation. An alternative model could be that H3 and H4 tails interact, in particular following acetylation of H4. This might sterically hinder docking of MOZ with H3. However, preliminary binding studies did not detect any interaction between core histones and unmodified or acetylated H3 and H4 peptides (data not shown).



**Figure 3.22. Quantitative analysis of MOZ DPF binding to site-specific histone H3 and H4 PTMs.** Fluorescent polarisation assay measuring the affinity of MOZ DPF binding to N-terminal FITC-labelled peptides including H3 unmodified peptide (1-21) (pink-circle), H4 unmodified peptide (2-24) (blue-square), H3 (1-21) acetylated at K9 (green triangle) and H4 (2-24) acetylated at K16 (orange diamond). The data are presented as averages and standard error of the means from triplicate samples.

### 3.12. Comparison of MOZ DPF to DPF3

Sequence alignments of the MOZ DPF domain revealed a high degree of sequence conservation (72% similarity) with the tandem PHD fingers of the human DPF family of proteins. Such high homologies suggest structural similarity and possible functional similarity with regard to binding preference for histone PTMs. The DPF3b PHD1&2 three-dimensional structure has recently been solved using NMR methods (Zeng et al., 2010). These structures were utilised to compare the conservation of surface exposed residues within the binding site. Protein Data Bank (PDB) NMR structures of DPF3b were manipulated in PyMOL™ (2006, DeLano Scientific LLC) by colour coding those amino acids conserved with the MOZ DPF. Light yellow and yellow coloured residues represent similar residues between DPF3b and MOZ PHD fingers 1 and 2 respectively and light orange and orange residues represent identical residues between DPF3b and MOZ PHD fingers 1 and 2 respectively (Figure 3.23). By mapping conserved surface binding residues onto the DPF3b structure, it is evident that almost the entire binding surface is composed of residues that are identical in MOZ (Figure 3.23). This suggested that DPF3b and MOZ interact with their histone peptide substrates in a similar way and additionally, leads to the assumption that the overall structure of the tandem PHD fingers is highly analogous. The DPF3b tandem PHD fingers establish contacts with the entire length of the histone H3 peptide (1-20), including the recognition of and contact to H3K9. This differs from other PHD fingers like the cognate BHC80 PHD domain, which contacts the first 8 residues of the H3 peptide with H3 residues K9 and S10 making numerous non-specific contacts with a neighbouring molecule (Lan et al., 2007). The

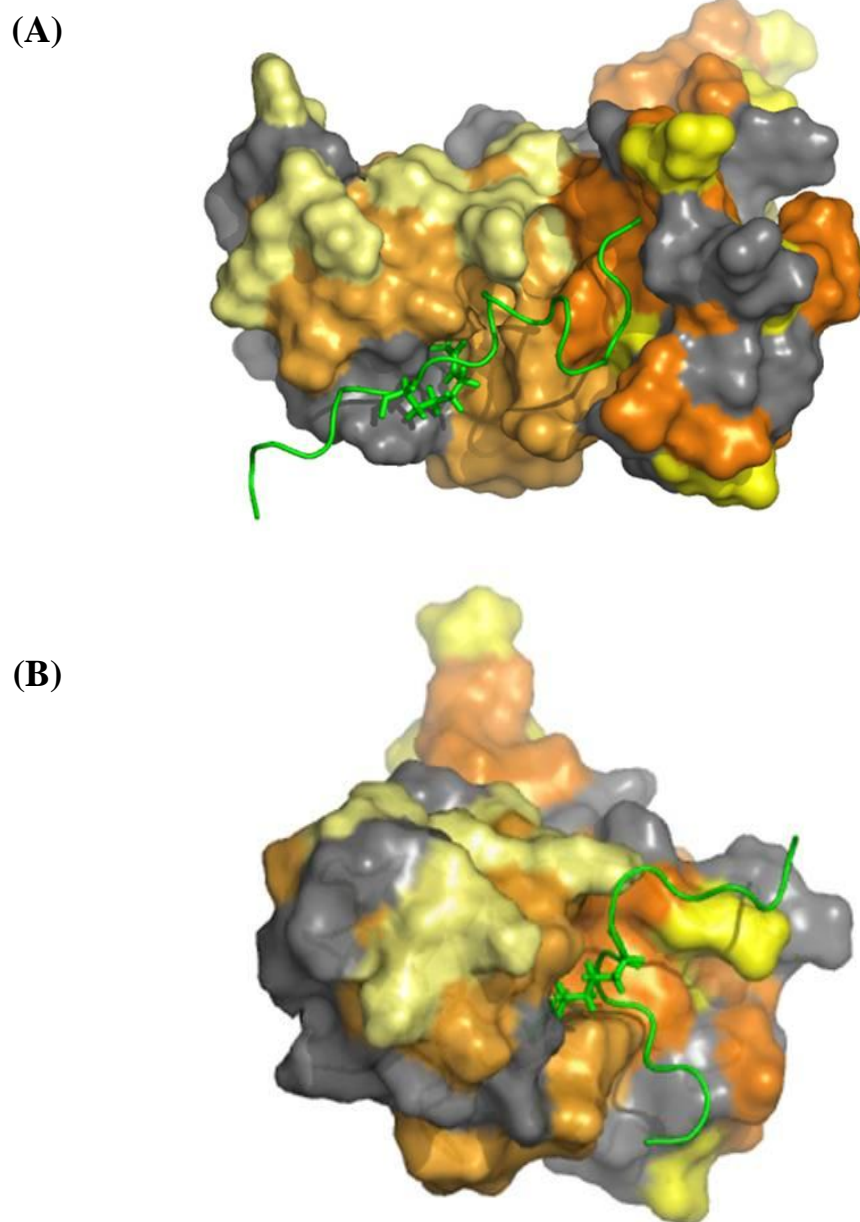
histone H3 peptide forms an extended conformation bound in an elongated groove of the double PHD finger (Figure 3.23(A)). DPF3b interaction with histone H3 is enhanced by acetylation at lysine 14 and inhibited by trimethylation at lysine 4. In comparison, the MOZ DPF was also shown to bind acetylated lysine residues of histone H3 and further, the MOZ DPF interaction with H3 is inhibited by trimethylation at H3K4. The acetylated H3K14 intercalates into a hydrophobic pocket of DPF3b PHD1, of which the key hydrophobic amino acid residues lining this pocket are all conserved in the MOZ sequence. In addition, those amino acids that secure the binding of the N-terminal H3 residues, including unmodified H3K4, are conserved. Interestingly, those residues located at the rim of the acetyl-lysine binding pocket, which are responsible for anchoring the peptide within the module, are not conserved in the MOZ sequence. This suggests that the MOZ DPF may adopt a similar mode of histone H3 recognition to DPF3b, but may utilise a slightly different mechanism to bind the acetyl-lysine PTMs.

The focal distinction between the two binding patterns is in the tolerance for H4 lysine acetylation. The DPF3b PHD1&2 preferentially binds acetylated H4 peptides, however with a 15-fold reduction in affinity compared to H3K14ac. The H4 peptide folds over itself extending the main chain peptide away from the binding site and allowing only the insertion of the K16ac group (Figure 3.23 (B)- side view). Therefore, a weaker interaction is seen compared to the H3K14ac peptide, due to fewer contacts being made with the entire length of the peptide. Thus, H4 acetylation enhances DPF3b binding to unmodified H4. In contrast, the acetylation of histone H4 prevents the association of MOZ DPF with histone H4. Referring more meticulously back to the sequence alignment of DPF3b and MOZ (Figure 3.3) shows that, despite a

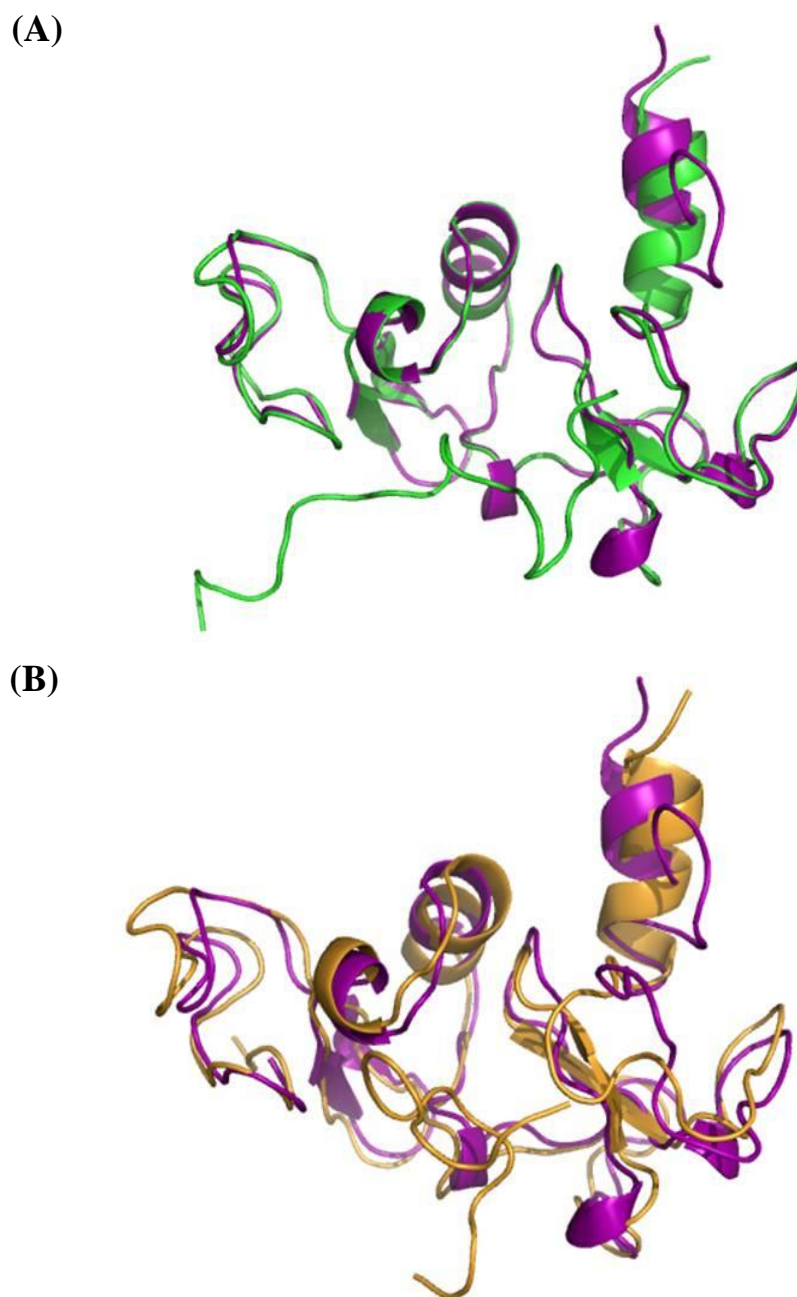
significantly high sequence homology, two residues in PHD1 and one residue in PHD2, which have been structurally implicated in DPF3b acetyl-H4 binding are not conserved in MOZ (Figure 3.3- **X**). This suggests that MOZ DPF adopts different H4 binding mechanisms, emphasising the importance of structural characterisation as small sequence differences may have a profound impact on functionality.

In addition to sequence alignments, structural homology modelling of the MOZ DPF was carried out using the Modbase web program designed by the Sali lab (University of California, San Francisco). The MOZ DPF sequence was entered into the database and two hits were returned; DPF3b (2KWJ) and MLL PHD3 (2KYU). Considering the previously observed high similarity to DPF3b, the predicted structure of MOZ DPF was modelled upon this databank entry (Figure 3.24- purple ribbon). In PyMOL<sup>TM</sup>, the predicted structure of MOZ DPF was superimposed with the structure of DPF3b in the presence of the H3K14ac (2KWJ) (Figure 3.24(A)- green ribbon) or H4K16ac (2KWN) (Figure 3.24(B)- orange ribbon) peptide. This figure predicts that the overall MOZ DPF structure favours the configuration of DPF3b in complex with an H3K14ac peptide and yet, in comparison, its consequent superposition onto the DPF3b-H4K16ac structure shows areas of misalignment especially within the loop secondary structure regions. These models are in agreement with the histone binding studies carried out in this chapter, suggesting that DPF3b and MOZ DPF may have differential abilities with regard to binding H4K16ac.





**Figure 3.23. Conservation of surface exposed residues based on the hDPF3 NMR structure.** (A) PDB; 2KWJ, DPF3b in complex with Histone H3K14ac peptide (green). (B) PDB; 2KWN, DPF3b in complex with Histone H4K16ac peptide (green). Light orange (PHD1) and orange (PHD2) colours represent residues that are identical between DPF3b and MOZ. Light yellow (PHD1) and yellow (PHD2) colours represent residues that constitute similar substitutions in both domains. The H3K14 and H4K16 residues are shown as stick representations. Models were created in PyMOL™.



**Figure 3.24. Alignment of DPF3b and the structural homology model of MOZ DPF.** Structural models were created by the Modbase web program designed by the Sali lab (University of California, San Francisco). Predicted structures of MOZ DPF (purple ribbons) superimposed on the DPF3b NMR structures; **(A)** the backbone structure of DPF3b (PDB 2KWJ) (green) from the DPF3b/H4K14ac peptide complex and **(B)** the structure of DPF3b (PDB 2KWN) (orange) from the DPF3b/ H4K16ac complex. Peptide structures are masked.

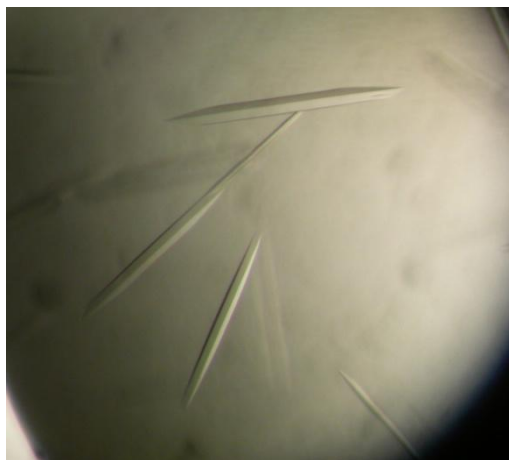
### 3.13. Structural characterisation of the MOZ DPF

Structural analysis is essential to determine the intricate details involved in the molecular recognition of specific histone PTMs by the MOZ DPF. To date, a collaboration has been set up with Dr. Ingrid Dreveny, a protein crystallographer with the Division of Medicinal Chemistry and Structural Biology at the University of Nottingham.

In conjunction with Dr. Baigong Yue, the expression and purification of His-tagged MOZ DPF recombinant protein was optimised to routinely yield a large quantity of soluble protein. Briefly, 5 L bacteria were cultured at 37 °C until the OD<sub>600nm</sub> reached 0.5-0.6. Protein production was initiated by the addition of 0.1 mM of the IPTG inducer and cultured for a further 16 hours at 22 °C. Cells were harvested, resuspended in binding buffer (0.02 M sodium phosphate, 0.5 M NaCl pH7.4), sonicated and the soluble protein was harvested by centrifugation. Supernatants were finally filtered through a 0.45 µm filter. Purification proceeded via a HiTrap Chelating column loaded with 50 mM NiSO<sub>4</sub>. The column was equilibrated with binding buffer, loaded with the protein sample and washed with binding buffer containing 5 mM imidazole. His-tagged MOZ DPF was then eluted from the column using binding buffer containing 300 mM imidazole and the His-tag cleaved from the protein by incubation with thrombin protease overnight at 22 °C. Approximately 60 mg of protein was purified from a 5 L starting volume of bacterial culture.

Recently, the structure of the MOZ DPF in complex with a histone H3 N-terminal peptide was solved to a high resolution of 1.7 Å. Co-crystallisation trials were set up by Dr. Ingrid Dreveny using MOZ DPF and a histone H3 N-terminal

peptide substrate. This yielded large, single crystals (Figure 3.25) that have the symmetry of space group  $P4_12_12_1$ . A high resolution dataset was collected at the European Synchrotron Radiation Facility, Grenoble, beam line ID23-2. The structure was solved by sulphur-SAD phasing using the anomalous signal from the zinc atoms. After density modification, model building was carried out and the structure was refined to a final crystallographic Rfactor of 17.3% and Rfree of 19.8%.

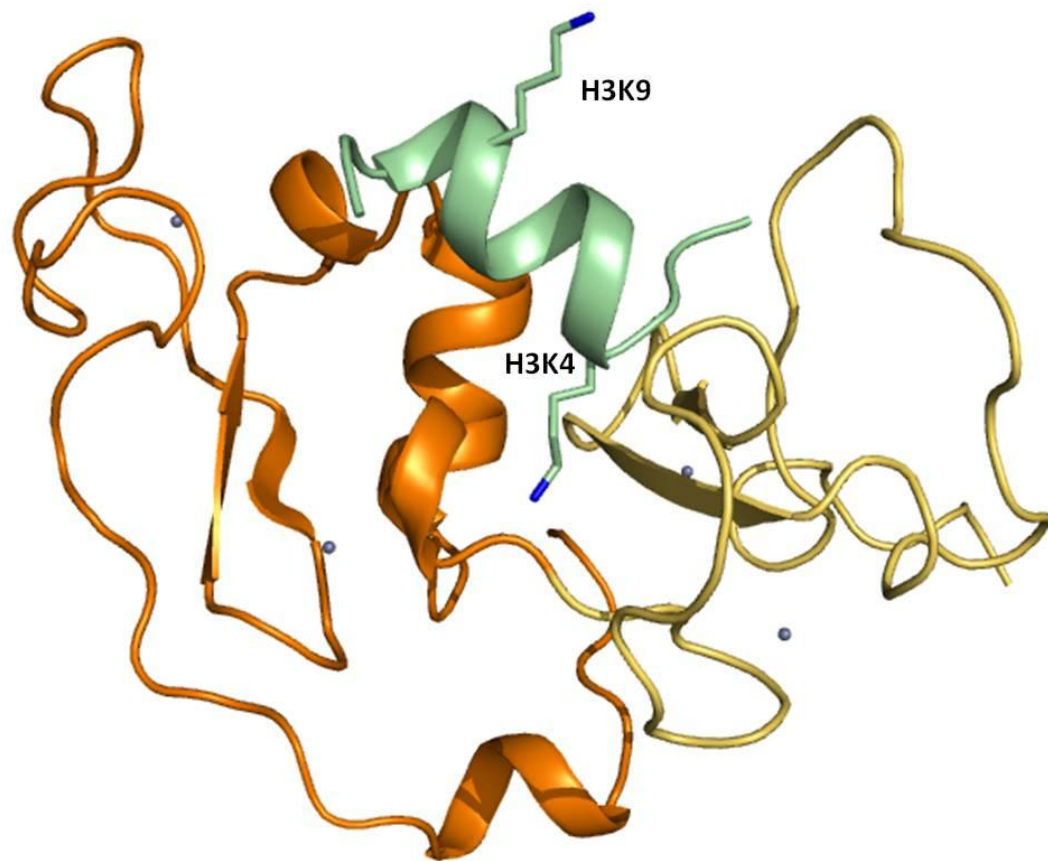


**Figure 3.25. Crystallisation of the MOZ DPF domain.** MOZ DPF was expressed in a bacterial system and purified via its His-tag. Purified protein was utilised in co-crystallisation trials with an N-terminal histone H3 peptide.

The structure is of very high quality and provides mechanistic insight into how both PHD1 and PHD2 contribute to engage the first thirteen N-terminal residues of H3. Unexpectedly, the H3 peptide (residues T3-G12) in the complex structure adopts a two-turn alpha-helical conformation (Figure 3.26), which to our knowledge has not been observed previously for the H3 N-terminus in interactions with PHD domains and is likely to be an induced fold. Further, the structure shows that the interaction between unmodified H3 and MOZ DPF is characterised by an intricate network of hydrogen bonding interactions. The assignment of DPF interactions with the entire length of the H3 peptide is still ongoing, however an example of how H3K4 is recognised is depicted in Figures 3.26-3.27. H3K4 appears to be embedded into a binding pocket constituted by amino acids from both PHD fingers 1 and 2 (3.26). Multiple hydrogen bonds are established between the free amino group of the H3K4 side chain and DPF main chain carbonyls (Figure 3.27). The crystal structure also highlights the existence of water-mediated hydrogen bonds stabilising the H3K4 side chain; contacts which cannot be detected using NMR solution structures (Figure 3.27). This is consistent with the biochemical data showing that PHD1 and PHD2 of the DPF composite functionally cooperate to sense the H3K4 modification status. Interestingly, H3K9 protrudes from the second  $\alpha$ -helical turn, away from the MOZ DPF binding site (Figure 3.26). Peptide binding assays showed binding of K9ac and K9me3 modified H3 peptides by MOZ DPF. The capture of this structure implies that modification at the H3K9 residue does not affect binding of the MOZ DPF to the H3 peptide, rather than specifically recognising the modification itself. However, it is possible that modification of H3K9 by methylation or acetylation could induce a change in the histone peptide conformational structure, to place the K9ac/me3 into

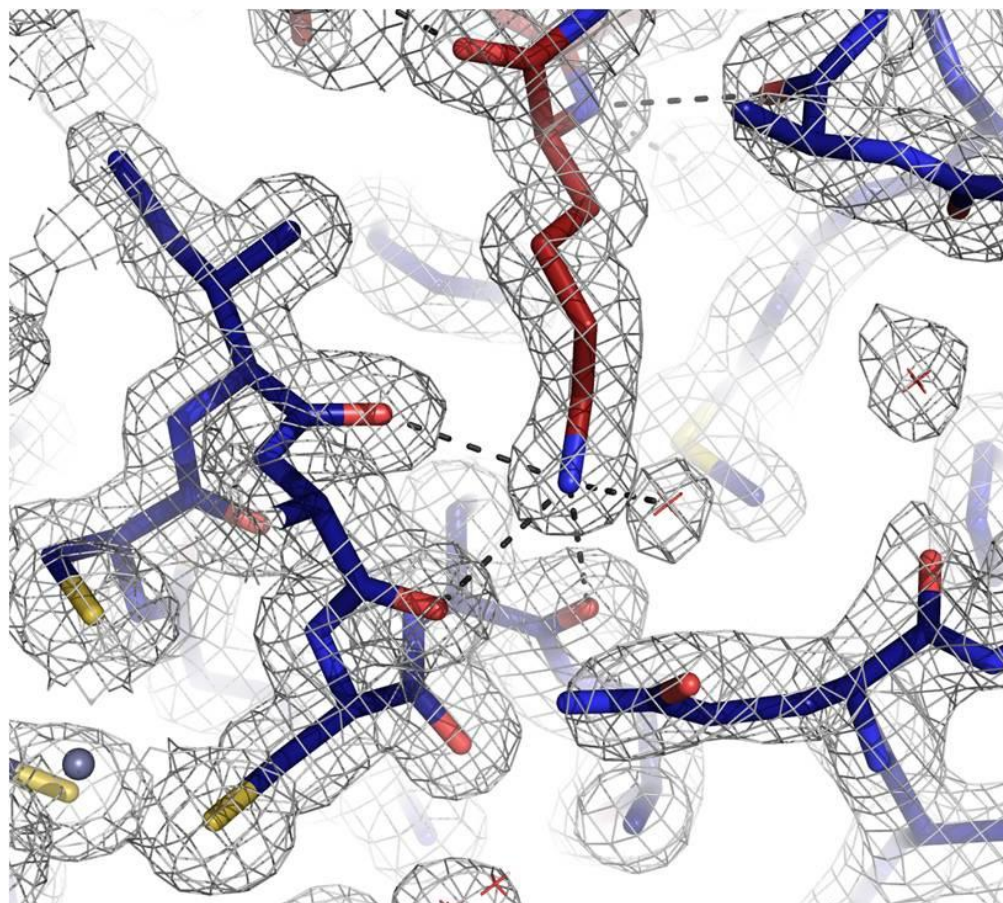
the MOZ DPF binding site. This has been seen previously for the H3 peptide in complex with DPF3b whereby, when unmodified, the K14 residue extends away from the binding site. However, upon acetylation the conformation of the peptide changes, projecting H3K14ac into a hydrophobic pocket created by the DPF3b tandem PHD fingers (Zeng et al., 2010). Co-crystallisation trials are currently being set up with MOZ DPF and an H3K9ac peptide.

Our MOZ DPF structure can be superimposed onto the DPF3b NMR structure (102 aligned residues, sequence identity of 52%) with a root mean square deviation value of 1.9 Å between the structures. The overall fold is conserved but differences are apparent in loop regions (Figure 3.28). However, the most striking difference between the complexes pertains to the conformation of the H3 peptides. In our structure the H3 peptide is  $\alpha$ -helical and more intimately embedded within the DPF binding site, whereas within the DPF3b complex the H3 peptide is unstructured (Figure 3.28).



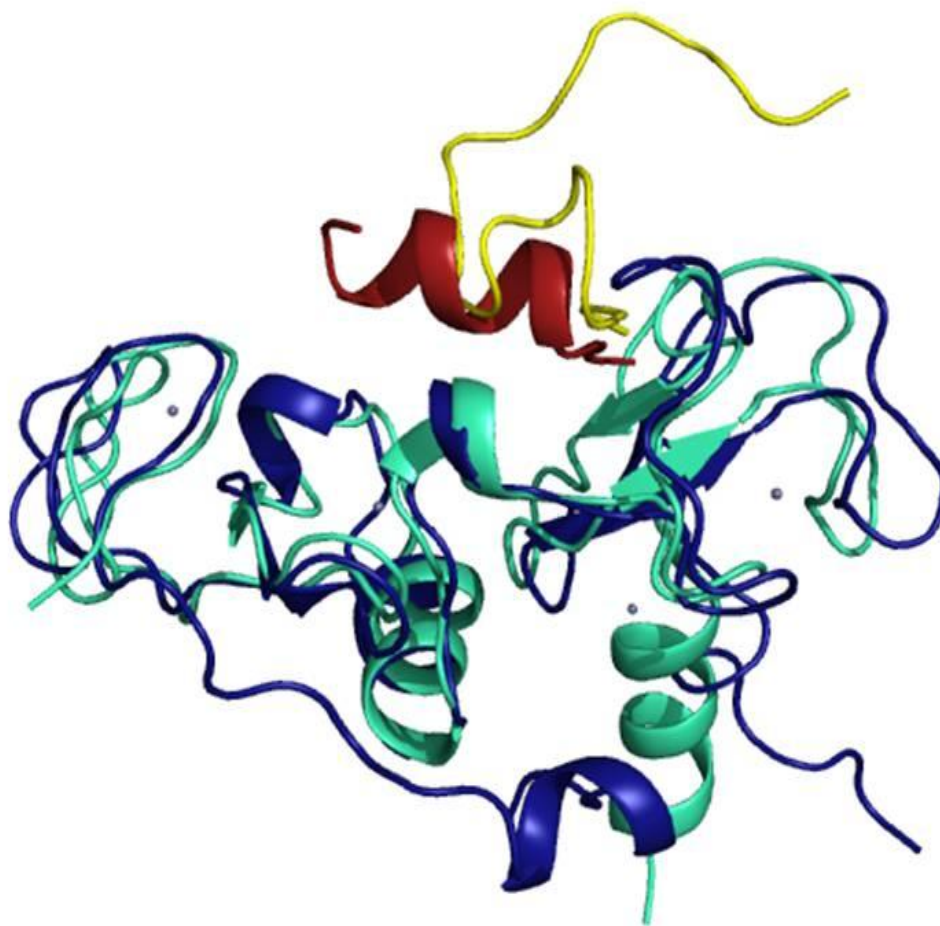
**Figure 3.26. Two-dimensional structure of the tandem PHD finger of human MOZ.** Cartoon representation of the MOZ DPF domain (PHD1 in orange and PHD2 in yellow) in complex with the unmodified histone H3 N-terminal peptide (green ribbon). Histone H3 lysine side chains, K4 and K9 are shown protruding into or away from the DPF binding site respectively. Coordinated zinc atoms are depicted in spacefill format as grey spheres.





**Figure 3.27. Structural mechanism of H3K4 engagement by human MOZ DPF.** A focussed view of the region surrounding H3 lysine 4 (red), detailing plausible hydrogen-bonding interactions (dashed lines) with DPF main chain carbonyls (blue) and a water molecule (red cross).





**Figure 3.28. Comparison of MOZ and DPF3b tandem PHD fingers.** Superposition of the MOZ DPF crystal structure (blue) with DPF3b (PDB 2KWK, (Zeng et al., 2010)) NMR tandem PHD finger structure (cyan) in complex with unmodified H3 N-terminal peptides (red and yellow respectively).

**3.14. Summary**

MOZ catalyses the acetylation of lysine residues within histone proteins (Champagne et al., 2001), however how MOZ is targeted to specific chromatin substrates in order to direct its HAT activity remains unclear. MOZ is a large modular protein and only the MYST domain has been functionally characterised, therefore at the outset of this work the functional importance of the N-terminal tandem PHD fingers was yet to be established. As the DPF domain is retained in oncogenic MOZ fusion proteins in AML (Katsumoto et al., 2008), it is likely to have important functions in both a normal and malignant context.

The PHD domain is a well-established versatile epigenome reader that can recognise and interpret the methylated (Li et al., 2006), acetylated (Lange et al., 2008) or the unmodified states (Lan et al., 2007) of histone lysine residues. MOZ and MORF are unusual in that they contain a double PHD finger arranged in tandem, but other chromatin modulators have also been shown to harbour double ‘reader’ domains. These include the TAF1 double bromodomains that bind multiple acetylated H4 peptides (Jacobson et al., 2000) and the two tudor domains of JMJD2A, that together cooperate to recognise the H3K4me3 modification (Huang et al., 2006). More recently, the structural and functional characterisation of histone recognition by the double PHD finger protein, DPF3b, has been established. The tandem PHD fingers were shown to function as an integrated unit that determines its recognition specificity for the histone PTMs (Zeng et al., 2010). Sequence alignments of the MOZ DPF domain with DPF3b revealed a high degree of sequence

conservation (53%) between the two sequences. Thus, we first asked whether the MOZ DPF domain could also function as a histone PTM sensor.

Using GST pulldown assays it was demonstrated that the N-terminus of MOZ (1-321) and the DPF alone bound strongly to histone H3 in a dose-dependent manner. However, only a very weak interaction was observed with histones H2B and H4. Mutagenesis of the DPF indicated that MOZ binding of histone H3 was indeed mediated by the DPF domain and that the integrity of both PHD fingers was required for the interaction.

Commercial core histones are purified from cell extracts and hence, exhibit a mixed population of PTMs. In order to define the specificity of MOZ DPF binding, we assessed how selected histone PTMs impact on these interactions. Using GST pulldown assays with biotinylated histone peptides as substrates, revealed that binding of MOZ DPF to histone H3 and H4 was mediated by the histone N-terminal tail sequences. Conversely, no interaction between the N-termini of H2A or H2B was observed. Further, pulldowns carried out with a series of modified H3 and H4 peptides, showed that both the MOZ N-terminus and DPF domain were unable to bind H3 containing trimethylation at K4. This result was also consistent with a low sequence similarity observed by the alignment of MOZ DPF with other H3K4me3-binding PHD fingers. Interestingly, MOZ 1-321 and the DPF alone were able to accommodate the dimethylation of H3K4, suggesting that the increased size and cationic radius of an additional methyl moiety is unfavourable for binding within the DPF binding pocket. Of course, to establish the actual effect of the third methyl group on binding requires the structural characterisation of the MOZ DPF in complex with the H3 peptide. Indeed, the latter cocrystallisation of MOZ DPF with an H3 peptide

rationalises the *in vitro* exclusion of H3K4me3 from the DPF binding site. In contrast to H3K4, trimethylation of H3K9 did not affect binding to either the MOZ N-terminus or the DPF domain. However, H3K9/14 acetylation was seen to have differential effects on the MOZ 1-321 and DPF constructs. H3Kac stimulated the interaction with MOZ 1-321, but appeared not to have an affect on the binding of DPF alone. This suggested that a H3 acetylated peptide may induce additional contacts with the MOZ 1-171 sequence. The MOZ 1-321 and DPF constructs both showed strong binding to histone H4, however no interaction was observed for both proteins to any acetylated form of the histone H4 peptide. This suggested that H4 acetylation was incompatible with MOZ recruitment to chromatin. Similar histone binding properties, namely the intolerance for H3K4me3 and acetylated H4, were also observed for the full-length MOZ protein and the closely related homolog MORF (87% identity). This implied that both enzymes are targeted to or released from chromatin enriched with similar PTMs.

Future work would entail an extensive profiling of the interactions between the MOZ DPF domain and other histone PTMs, not only to include individual PTM sites but also, to examine the effects of neighbouring modifications on recognition and binding. Recently, this has been made possible by the availability of commercial arrays such as MODified<sup>TM</sup> Histone Peptide Arrays (Active Motif). Each array contains 384 different histone modification combinations for the acetylation, methylation, phosphorylation and citrullination on the N-terminal tails of all four core histones; and thus would allow the rapid profiling of MOZ DPF interaction with multiple histone PTMs. In addition, recent data from our group shows that, like TIP60, MOZ expression is upregulated during DNA damage (unpublished). For this

reason, we would also like to investigate whether the MOZ DPF can function in sensing damage-associated PTMs, such as  $\gamma$ H2A.X or H3K56ac.

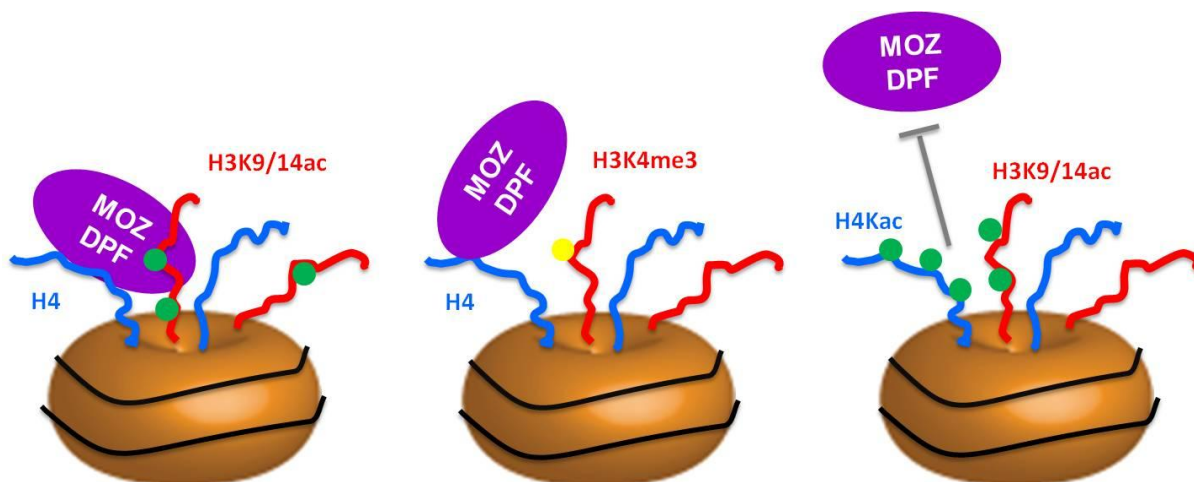
We next assessed whether the individual PHD fingers performed specific roles in governing the DPF binding pattern. Using GST pulldowns with constructs PHD1 (194-263) and PHD2 (250-323) and biotinylated histone peptides it was observed that PHD1 was able to bind unmodified H4 only, whereas PHD2 could bind both unmodified H3 and H4. Remarkably, PHD2 showed robust interactions with all modified N-terminal H3 peptides, including the trimethylated H3K4 mark. Noting the intolerance of the MOZ DPF protein for H3K4me3 binding, this result indicates that PHD1 and PHD2 functionally cooperate to determine the histone PTM binding selectivity of the DPF, in particular with regards to the discrimination of H3K4me3. Thus, advocating that PHD1 and PHD2 functionally cooperate to determine the histone PTM binding selectivity of the DPF, in particular to the sensitivity to H3K4me3. To establish functional cooperation, structural data would need to be obtained, to check that both PHD fingers make contacts with the histone peptide substrate. In fact, the latter cocrystallisation of the MOZ DPF and an H3 peptide confirms that residues from both PHD1 and PHD2 engage the N-terminal histone H3 peptide.

ChIP analysis and immunofluorescence revealed that MOZ could associate with chromatin enriched with H3Kac PTMs and, in part, with trimethylated H3K9. These data corroborate other *in vivo* studies showing that MOZ is required for the acetylation of H3K9 at transcriptionally active *Hox* gene loci (Voss et al., 2009). Conversely, MOZ was not immunoprecipitated or colocalised with chromatin containing the H3K4me3 modification. MOZ has been shown to activate certain

genes, including *Hox* genes (Voss et al., 2009), however our results showed that MOZ is excluded from chromatin regions that contain H3K4me3; a modification generally associated with transcriptional activation. However, it is not clear whether all transcribed genes contain this PTM, so our results suggested that the function of MOZ in gene regulation either precedes or comes after the deposition of H3K4 trimethylation. These *in vivo* data are consistent with our *in vitro* binding studies. It would be interesting to examine the occupation of genomic sites associated with MOZ complexes, once better ChIP-grade antibodies specific to MOZ become available. Using the ChIP-Seq technique MOZ-tagged binding sites can be correlated with one or more chromatin PTMs to verify our results in a promoter context. In addition, after the identification of MOZ binding sites, the affects of MOZ deficiency on histone PTMs at specific gene regions could be assessed by histone PTM-specific ChIP followed by quantitative genomic PCR.

The coexistence of multiple modifications on one or more histone tails of the same nucleosome suggests that chromatin regulators can sense multiple PTMs cooperatively or simultaneously. Evidence for the combinatorial recognition of multiple PTMs is emerging, for example the mouse TAF1 homologue, Brdt, which cooperatively binds two acetylation marks on a single H4 tail (Moriniere et al., 2009). Thus, we investigated as to how combinations of H3 and H4 PTMs influence the engagement of MOZ DPF. Using GST pulldowns with equimolar mixtures of peptides, it was demonstrated that the ability of MOZ DPF to bind H3, H3K9/14ac and H4 peptides was unaffected by H3K4me3, but inhibited by mono- or hyper-acetylation of H4. These results suggested a model where H3K4me3 acts as a recessive exit signal that does not affect the simultaneous binding to PTM residues on

neighbouring histone tails, whereas H4Kac exerts a dominant inhibitory signalling effect to prevent the binding of nearby modified histone tails by MOZ DPF (Figure 3.29).



**Figure 3.29. Schematic model depicting how combinations of H3 and H4 PTMs influence the engagement of MOZ DPF to chromatin.**

Acetylation of H4 blocks the interaction of MOZ DPF with other histone PTMs. How this occurs is puzzling, given that *in vitro* binding assays demonstrated that binding of H4Kac peptides to the DPF is below the detection limit of the assay. Using fluorescence polarisation anisotropy revealed that GST-MOZ DPF was able to bind a FITC-labelled H4K16ac peptide, albeit with a much lower affinity compared to H4. There are a number of possible explanations for the dominant inhibitory affect observed for H4Kac with regards to the binding of other histone PTMs. Firstly, the interaction with H4 may be transient, but sufficient to reduce binding of H3 peptides to DPF; it is possible that even a weak affinity for an acetylated H4 tail might induce

a conformational change to displace the H3 tail, if the two peptides bind at two distinct pockets within the DPF. The MOZ DPF was also demonstrated to bind histone H3 with a substantial increase in affinity compared to H4 and thus, given this apparent difference in binding strengths, displacement of H3 through competitive binding at one binding pocket seems a less likely explanation. It will be necessary to generate a crystal structure for DPF/H4 complexes to establish whether H3 and H4 N-terminal tails dock at similar or distinct site(s) on MOZ DPF. Finally, H3 and H4 N-terminal tails might interact with each other in a PTM dependent manner; however, preliminary experiments to detect H3 and H4 interactions failed to support this hypothesis.

Binding studies thus far have been conducted with histone peptide surrogates, however little is currently known how the combinations of histone PTMs are interpreted at the level of nucleosomes; a more physiologically relevant substrate. Recently, it has been reported that histone marks constitute unique nucleosomal patterning, which dictates critical chromatin associations (Ruthenburg et al., 2011). It would therefore, be important to assess the binding preference of MOZ DPF for histone PTMs at the nucleosomal level. Nucleosomal substrates could be obtained via two different methods; firstly, the reconstitution of semi-synthetic histones, bearing the desired combinations of PTMs, into octamers. Alternatively, by the preparation of mono- or oligo-nucleosomes from native chromatin extracts by MNase digestion and sucrose gradient ultracentrifugation. These substrates would then be used to assess the enrichment or exclusion of the selected nucleosomal PTMs with or from MOZ.

The cocrystallisation of MOZ DPF with an N-terminal histone H3 peptide has produced a high resolution structure that provides mechanistic insight into how both

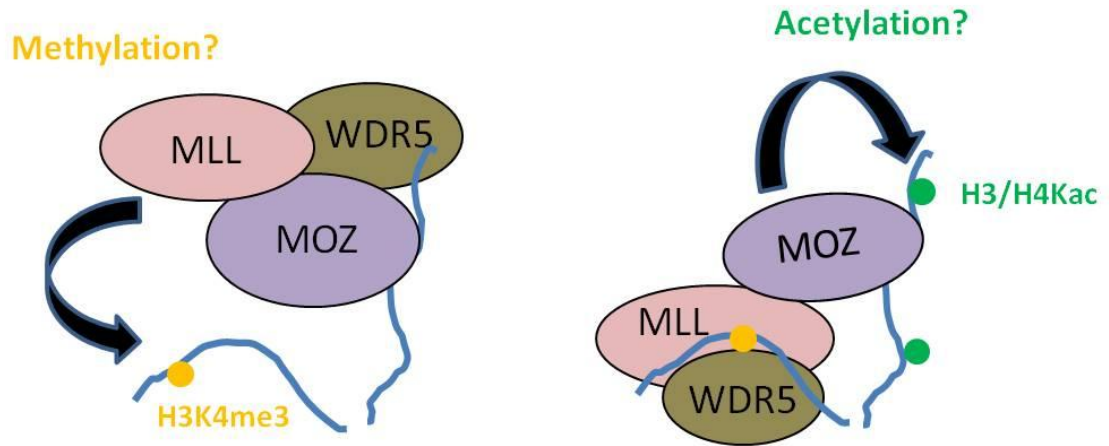


PHD1 and PHD2 contribute to engage the first thirteen N-terminal residues of H3. The assignment of DPF interactions with the entire length of the H3 peptide is still ongoing, however interactions mediating H3K4 recognition have been defined. H3K4 was inserted into a binding pocket constituted by PHD1 and PHD2 residues and further, shown to be stabilised by multiple hydrogen bonds involving DPF main chain carbonyls and a water molecule. These results rationalised the exclusion of H3K4me3 from the MOZ DPF binding site and are consistent with biochemical binding studies, whereby both PHD fingers of the DPF composite cooperate to sense the H3K4 modification status. Remarkably, the complex structure revealed that the H3 peptide adopted an  $\alpha$ -helical structure, which to our knowledge has not yet been observed for the H3 N-terminus in interactions with PHD domains. Future work will entail the cocrystallisation of the MOZ DPF with H3K9ac, in order to determine the effect of the H3K9 modification on the fold of the H3 peptide. We also wish to characterise the DPF/H4 complex to examine whether the H3 and H4 peptides bind in similar or two distinct binding site(s). Each of the above structures would also be useful to guide the mutagenesis of DPF residues involved in sensing permissive and non-permissive PTMs. Binding studies would then be carried out with such mutants, to examine how the structure-function influences the MOZ DPF engagement with histones.

In summary, our data shows that the MOZ DPF allows it to engage chromatin that is enriched in H3Kac but that lacks H3K4me3 and acetylation at H4. The binding preference of MOZ DPF has been established however, the order of PTM binding is not clear. Thus, to provide insight into MOZ function and the sequence of events that takes place, further probing is required into the complexes in which MOZ resides. For

instance the MOZ/EAF6/BRPF/ING5 complex harbours multiple chromatin ‘reader’ modules each with distinct, and in some cases, opposing, binding specificities. The trimethylation of H3K4 is bound by the ING5 PHD finger (Champagne et al., 2008), but is excluded from the MOZ DPF. It has recently been shown that the combinatorial action of multiple PHD domains is essential in regulating the chromatin binding and substrate specificity of HBO1 HAT-containing complexes (Saksouk et al., 2009). Thus, it would also be important to analyse MOZ DPF binding function in the context of affinity purified complexes.

MOZ has also been found in complexes containing the methyltransferase MLL and the adaptor protein WDR5, which catalyses the trimethylation of H3K4 (Paggetti et al., 2010). Such complexes provide a link between histone acetylation and methylation events at gene promoters. Recruitment of MLL and WDR5 by MOZ may then drive the deposition of H3K4me3, which is incompatible with MOZ and thus, serves as a release signal from chromatin, unless a free H4 tail is available. Conversely, MOZ could interact with chromatin bound MLL/WDR5, which facilitates MOZ acetylation of the same or neighbouring histone tails (Figure 3.30). These recruitment/release events might allow MOZ complexes to propagate selected PTMs in chromatin. Interrogation of such models will begin to shed light on the molecular mechanisms by which MOZ functions.



**Figure 3.30. Schematic model depicting the possible sequence of events involved in MOZ/MLL/WDR5 complex histone acetylation and methylation.**

## **Chapter Four: Results**

**The MOZ DPF domain exhibits a H3-specific  
acetyltransferase activity**

### 4.1. Introduction

The previous chapter has demonstrated that the MOZ DPF functions to engage chromatin by binding the histone tails of H3 and H4. The two PHD fingers cooperate to determine selective interactions with histone PTMs. Intriguingly, MOZ appears unable to make a stable association with the acetylated form of histone H4. This is of particular interest due to the presence of an adjacent MYST domain in MOZ, which contains an *in vitro* acetyltransferase activity shown to target H4 acetylation (Champagne et al., 2001) and more recently, the acetylation of non-histone substrates (Ullah et al., 2008) (Collins and Heery, unpublished data). Thus MOZ may regulate its own recruitment to and exit from chromatin. A conundrum in the field is that knockout studies revealed that MOZ ablation leads to reduced H3K9 acetylation, whereas H4 acetylation is unaffected (Voss et al., 2009). This chapter describes work carried out to investigate the substrate specificity of MOZ and whether the DPF influences the target of MOZ acetylation.

In addition to a histone PTM ‘writing’ capacity, the majority of chromatin modifying enzymes also contain a ‘reader’ module and recent studies show that both these activities can cooperate in synergy to obtain the overall desired function. For example, the bromodomain of p300 is required for stimulating p300 HAT activity in order to maintain the basal level of histone acetylation and to induce transcriptional activation of p300-dependent genes (Chen et al., 2010). Further, the bromodomain is important for p300 to function at full coactivator activity, whereby the bromodomain recognises acetylated lysine residues and stabilises the association of p300 to specific chromatin sites (Zeng et al., 2008b). Examples of functional synergy between

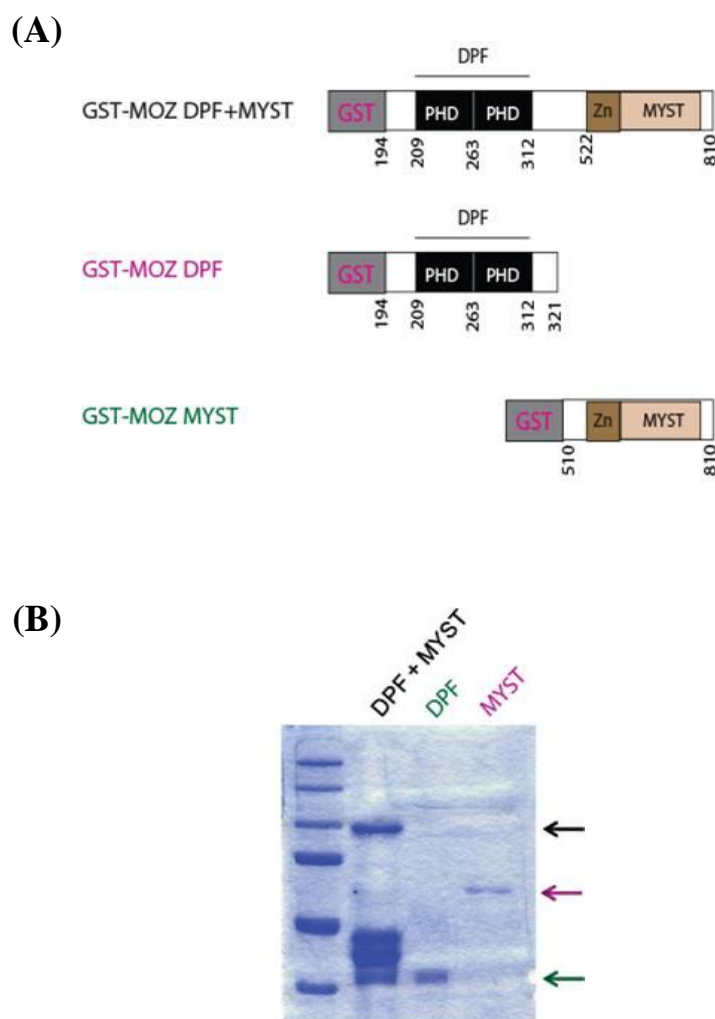
‘reader’ and ‘writer’ modules within the same protein are emerging and so, this chapter also describes work carried out to test whether the DPF module affects other functional properties of the MOZ protein; including its acetyltransferase and transcriptional coactivator activities and its subcellular localisation.

### 4.2. The MOZ DPF contains an intrinsic HAT activity

To investigate whether the DPF influences the histone acetyltransferase activity of MOZ, three constructs incorporating the individual DPF and MYST domains alone and the DPF+MYST domains together (Figure 4.1(A)), were cloned into the bacterial GST-expression plasmid. GST-fusion proteins were expressed in the Rosetta *E. coli* strain and purified via the high affinity of GST for glutathione using glutathione cross-linked sepharose beads. Proteins were eluted off the beads and their concentrations normalised by SDS-PAGE and Coomassie staining to ensure equal amounts of protein were used in subsequent *in vitro* assays (Figure 4.1(B)). The Coomassie gel highlighted normalised concentrations of GST-MOZ DPF and GST-MOZ DPF+MYST proteins, however it also showed an unequal concentration of GST-MOZ MYST in comparison (Figure 4.1(B)- compare lanes 2&3 with 4). To rectify this, five times the amount of GST-MOZ MYST protein loaded onto the gel was used in the following HAT assays.

The constructs GST-MOZ MYST, GST-MOZ DPF+MYST and the control construct GST-MOZ DPF were then utilised in both filter binding and in-gel HAT assays. Firstly, filter binding assays, as described in the Materials and Methods, were carried out to measure the total acetylation of substrate, where fusion proteins were

incubated with core histone extracts as a substrate and [ $^3\text{H}$ ]-Acetyl-CoA as the acetyl group donor. Samples were incubated at 30 °C for 30 minutes to allow the reaction to reach completion and spotted onto filter paper. Filters were washed thoroughly with sodium carbonate solution to remove residual radioactive isotope and incubated in scintillation fluid overnight. Tritium counts were measured using a scintillation counter and results calculated relative to a background reading carried out without HAT enzyme or with the GST tag protein alone.



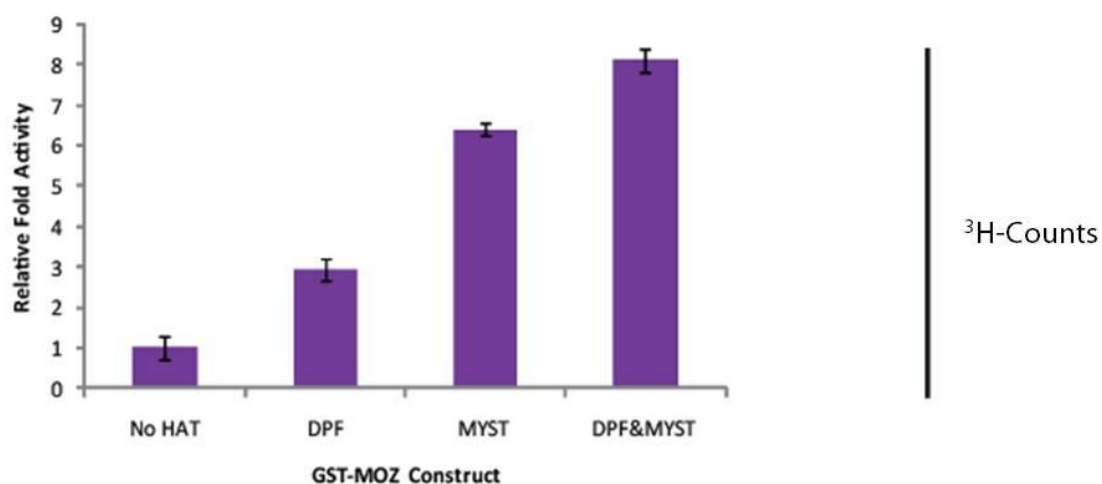
**Figure 4.1. Purification of the MOZ DPF and MYST domains.** (A) Schematic representation of GST-MOZ fusion constructs used in subsequent assays. (B) Normalisation of GST-MOZ concentration by SDS-PAGE and Coomassie staining.



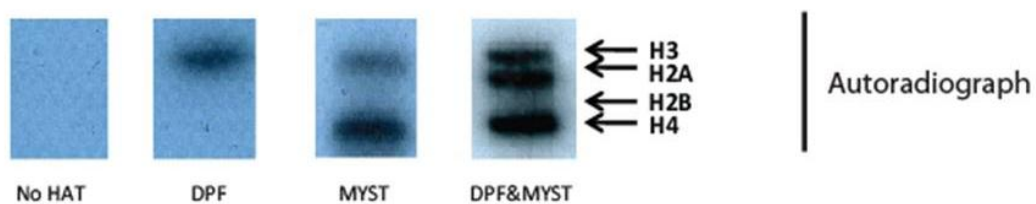
As shown in Figure 4.2, the DPF+MYST exhibited an 8-fold acetylation activity on substrate histones compared to a 6-fold activity for the MYST domain alone (Figure 4.2(A)- compare bars 3&4 to 1). This suggests that in the presence of the DPF the MYST domain acetyltransferase activity is slightly enhanced, perhaps due to the gain of a histone docking module. Surprisingly, the DPF alone, included in the assay as a negative control, exhibited an acetylation activity of histones that was 3-fold over background (Figure 4.2(A)- bar 2).

To investigate this further, an in-gel HAT assay was conducted to obtain data on the specific histone target of enzyme acetylation. Amendments to the above protocol included the use of [ $^{14}\text{C}$ ]-Acetyl-CoA radioisotope as the acetyl group donor and following the 30 minute incubation, reactions were subjected to SDS-PAGE and autoradiography. As shown in Figure 4.2, GST-MYST and GST-DPF+MYST both showed an ability to acetylate H4. However, the DPF+MYST substrates also included the acetylation of H3 and possibly H2A. In contrast, the control showed no detectable acetylation of core histones (Figure 4.2 (B)). This also shows that the histone preparations did not contain measurable levels of contaminant HATs. Remarkably, DPF was found to acetylate histone H3 but not other histones (Figure 4.2 (B)). This result suggests that the DPF appears to have a H3-specific HAT activity and its presence with the MYST domain extends the range of histone substrates for MOZ. This represents, to our knowledge, the first report of a mild acetyltransferase activity described for a PHD domain.

(A)



(B)



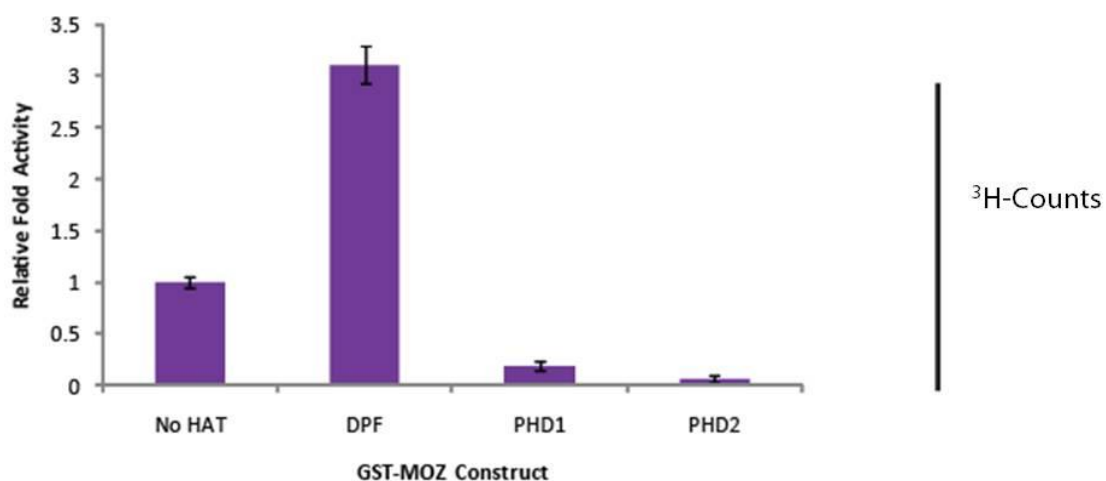
**Figure 4.2. The MOZ DPF contains an intrinsic HAT activity.** *In vitro* HAT assays using 1.5  $\mu$ g of HAT enzyme, 10  $\mu$ g core histones as substrate and either (A) [ $^3$ H]-Acetyl-CoA or (B) [ $^{14}$ C]-Acetyl-CoA as acetyl group donors. HAT reactions were followed by either (A) filter binding and liquid scintillation counting or (B) SDS-PAGE and autoradiograph analysis. Error bars represent the average of a minimum of two data sets.

To investigate the nature of the intrinsic HAT activity of the DPF module, HAT assays were performed with MOZ-DPF, MOZ-PHD1 and MOZ-PHD2 constructs. Both in-gel autoradiograph and filter binding tritium counts confirm previous results whereby the MOZ-DPF shows a 3-fold increase in acetylation activity towards histone H3 (Figure 4.3(A)(B)- lane 2). However, neither PHD1 nor PHD2 individually showed any HAT activity in these assays (Figure 4.3(A)(B)- compare lane 2 to 3 & 4). This result indicates that the acetyltransferase activity is an intrinsic property of the intact DPF module, as disrupting the tandem PHD finger structure obliterates activity. Comparable to the definition of DPF histone PTM binding selectivity, the recent results suggest that PHD1 and PHD2 also functionally cooperate to govern acetyltransferase activity.

In order to verify the specificity of histone acetylation activity for the full-length MOZ protein alternative expression methods were required, as the MOZ protein was too large to express efficiently as a GST-MOZ fusion. Therefore, the full-length protein and construct MOZ 1-810 (H1/5+DPF+MYST) were over-expressed as FLAG-tagged constructs in HEK293 cells. After 48 hours post-transfection, cells were lysed and the proteins immunoprecipitated with anti-MOZ (127A) antibody. This antibody was produced by Leukaemia and Lymphoma Research and purified by Dr. Hilary Collins. The MOZ 127A antibody is raised against the MOZ 172-321 epitope. IP samples were then used in an in-gel HAT assay. As shown in Figure 4.4, both FLAG-MOZ constructs were successfully immunoprecipitated, however it should be noted that the IP of the full-length protein was much weaker (Figure 4.4(A)- compare lanes 4&6). These results indicate that the MOZ protein contains two distinct acetyltransferase activities; the DPF which appears to selectively target

H3 and the MYST domain for which H4 is the major target. Together as DPF+MYST or in the context of the 1-810 protein, these domains appear to acetylate both H3 and H4, although 1-810 shows a preference to acetylate H3. How this substrate preference is established and what the lysine targets are, remain to be established. Moreover, whether histone PTMs affect these activities is also of great interest. The full-length MOZ also appears to target H3 and H4 in *in vitro* assays, consistent with our findings. However, these activities may be modulated by other proteins *in vivo*, such as suggested for the ING5 complex (Doyon et al., 2006).

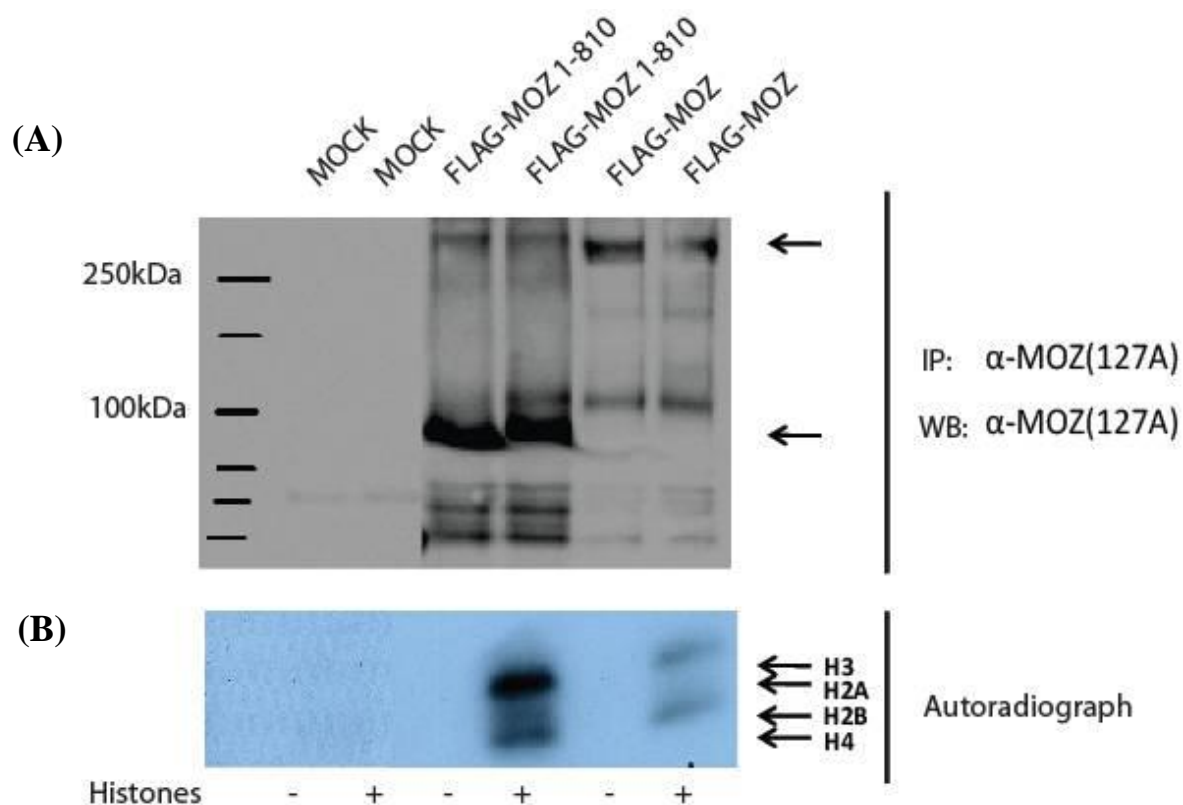
(A)



(B)



**Figure 4.3. HAT activity is a property of the intact DPF module.** *In vitro* HAT assays using 1.5  $\mu\text{g}$  HAT enzyme, 10  $\mu\text{g}$  of core histones as substrate and either (A) [ $^3\text{H}$ ]-Acetyl-CoA or (B) [ $^{14}\text{C}$ ]-Acetyl-CoA as acetyl group donors. HAT reactions were followed by either (A) filter binding and liquid scintillation counting or (B) SDS-PAGE and autoradiograph analysis. Error bars represent the average of a minimum of two data sets.



**Figure 4.4. Acetylation activity of the full-length MOZ protein.** HEK293 cells were mock transfected or transiently transfected with expression plasmids encoding FLAG-MOZ. After 48 hours post-transfection, cells were lysed and the proteins were immunoprecipitated with anti-MOZ 127A antibody. IP constructs were used as enzymes in *in vitro* HAT assays with 10  $\mu$ g core histones as substrate and [ $^{14}$ C]-Acetyl-CoA as an acetyl group donor. HAT reactions were followed by SDS-PAGE and **(A)** immunodetection with  $\alpha$ -MOZ 127A or **(B)** autoradiograph analysis.

### 4.3. Trimethylation of histone H3K4 blocks acetylation of histone H3

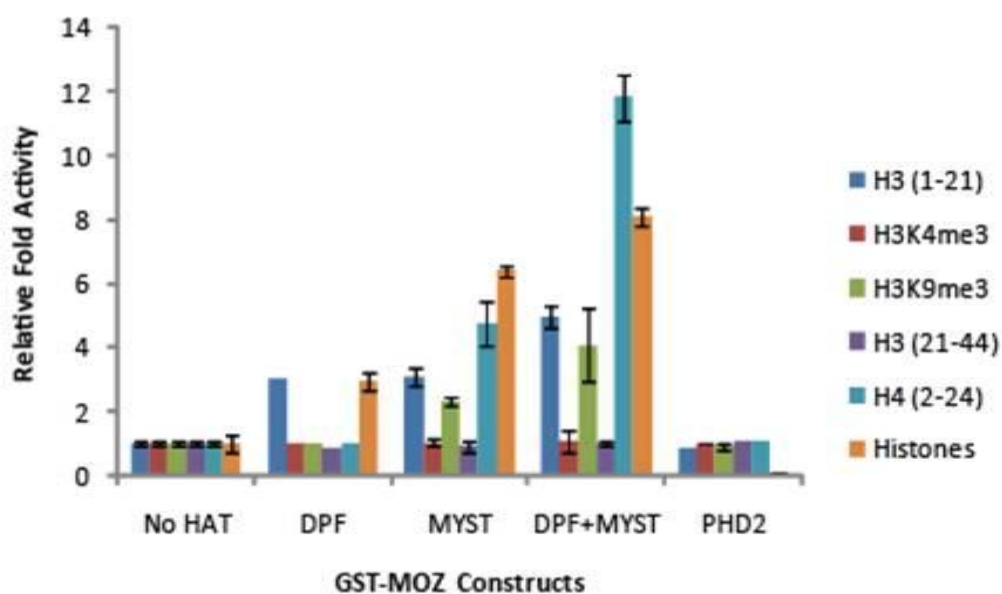
MOZ contains two distinct acetyltransferase activities that may cooperate to establish histone acetylation events in chromatin. As the DPF appears to have a dual function; as a H3/H4 PTM sensor and an acetyltransferase this raised the question regarding how existing PTMs on histones might affect the HAT activity. In other words we were interested to determine whether PTMs that promote or disrupt MOZ DPF binding to histones would influence the ability of MOZ to acetylate H3 and H4 N-terminal tails.

In order to investigate the effect of histone PTMs on MOZ HAT activity, *in vitro* HAT assays (filter) were conducted as previously described, using unmodified and modified histone peptides as substrates. We focussed on the following range of PTMs that included both modifications that promoted and those that disrupted DPF binding to the histone tails. All data was expressed as a fold activity relative to a background reading achieved without HAT enzyme. The GST-MOZ PHD2 protein was used as an assay negative control due to its strong histone binding capability but inability to acetylate histones.

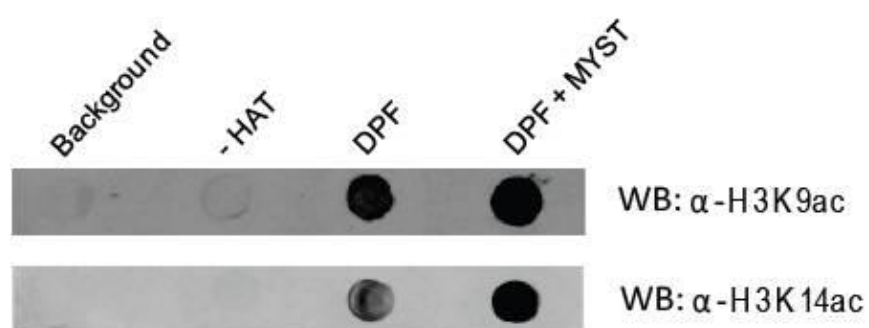
Both no HAT and PHD2 controls detected no acetylation of any of the five peptides or core histones. DPF alone was able to acetylate core histones and the H3 (1-21) peptide (Figure 4.5, orange and dark blue bars), but not the histone H3 (21-44) peptide (Figure 4.5, purple bar). This suggests that the target for MOZ DPF acetylation is located at the N-terminus of histone H3 at either lysine 9, 14 or 18, rather than in internal H3 sequences. No acetylation of the H3 peptide was observed

when modified at H3K4me3 or H3K9me3 (Figure 4.5 red and green bars). This result can be explained in several ways. We have previously shown that MOZ DPF does not bind trimethylated H3K4, hence it may not be able to acetylate the peptide. DPF can however bind to trimethylated H3K9 and so this indicates that K9 may be a major target of MOZ acetylation, thus no additional acetylation is possible. This suggests that K14 or other lysines may not be the target of MOZ DPF acetylation. In order to support this finding, HAT assays were carried out with the GST-MOZ DPF protein, unmodified histone H3 peptide substrate and the [<sup>3</sup>H]-Acetyl-CoA acetyl group donor. HAT reactions were then spotted onto nitrocellulose and subjected to immunoblotting with  $\alpha$ -H3K9ac and  $\alpha$ -H3K14ac antibodies. Control reactions show no background acetylation of histones underscoring the validity of the assay (Figure 4.6- lanes 1&2). MOZ DPF+MYST appears to acetylate both lysine residues 9 and 14 (Figure 4.6- lane 4, top & bottom panel respectively), however in comparison the MOZ DPF domain showed a preference for acetylating lysine 9 over 14 (Figure 4.6- lane 3, compare the spot intensity of top & bottom panel). These results suggest a HAT activity targeting H3K9 for the MOZ DPF.





**Figure 4.5. Trimethylation of histone H3K4 blocks the acetylation of histone H3.** *In vitro* HAT assays using 1.5  $\mu\text{g}$  HAT enzyme, 0.75  $\mu\text{g}$ /12.5  $\mu\text{M}$  histone peptides as substrate and [ $^3\text{H}$ ]-Acetyl-CoA as an acetyl group donor. HAT reactions were followed by filter binding and liquid scintillation counting. Error bars represent the average of a minimum of two data sets.



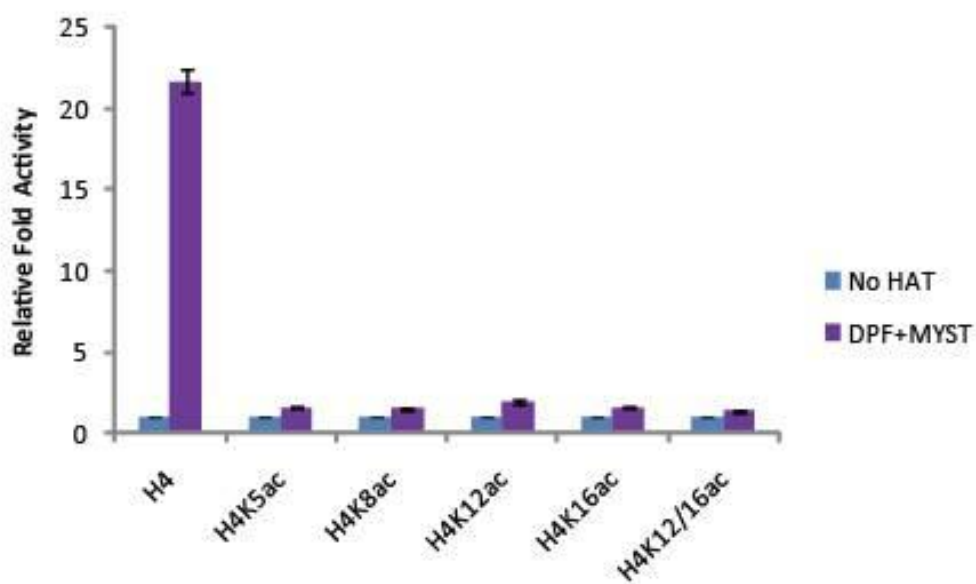
**Figure 4.6. MOZ DPF primarily acetylates lysine 9 of histone H3.** *In vitro* HAT assays using 1.5  $\mu\text{g}$  HAT enzyme, 0.75  $\mu\text{g}$ /12.5  $\mu\text{M}$  of unmodified H3 peptide as substrate and [ $^3\text{H}$ ]-Acetyl-CoA as an acetyl group donor. HAT reactions were spotted onto nitrocellulose membrane and followed by immunodetection of the indicated post-translational modifications.

The MYST domain was able to acetylate histone H4 to a greater extent than H3 (1-21), but not H3 (21-44) in the context of surrogate peptide substrates (Figure 4.5- light blue, blue, purple bars respectively). No acetylation of H3 was detected with a H3K4me3 peptide, although some acetylation of H3K9me3 was observed (Figure 4.5- red and green bars). The combined DPF+MYST construct showed strong acetylation of H3 and H4 which was blocked by H3K4me3 (Figure 4.5). H3K9me3 modified peptide only slightly decreased the ability of MOZ MYST and DPF+MYST to acetylate the H3 peptide (Figure 4.5- green bar). This suggests that the MOZ MYST domain not only acetylates H3K9, accountable for the slight decrease in acetylation counts, but also other lysine residues within the H3 peptide as acetylation is not completely abolished. The specific lysine target of MOZ acetylation still remains elusive, however this data implies that rather than searching for a single target, MOZ could acetylate multiple lysine residues within the H3 peptide. Indeed, HAT reactions followed by dot blot immunodetection suggest both H3K9 and H3K14 as targets for MOZ-DPF+MYST acetylation (Figure 4.6).

Remarkably, not one of the MOZ DPF, MYST or DPF+MYST proteins was able to catalyse the acetylation of a H3 peptide tri-methylated at lysine 4, despite all target lysine residues being free for modification (Figure 4.5- red bar). The lack of acetylation displayed by the MOZ HAT activities could possibly be a result of two mechanisms. Firstly, previous binding assay results show that the MOZ DPF exhibits an intolerance for the H3K4me3 modification. This finding suggests that MOZ is unable to dock onto peptides containing the H3K4me3 mark, preventing further acetylation of lysine residues on the same histone tail. Secondly, it was seen in the peptide acetylation assays that the MOZ MYST domain was also unable to acetylate

the H3K4me3 peptide. This implies that modification at H3K4 could possibly prevent or affect the entry of the substrate into the MYST catalytic active site, either by size or charge exclusion or a change to the conformation of the histone peptide, impeding acetylation of the histone lysine substrate.

Again, utilising peptides in the HAT assay enables the prediction of several possible lysine residues (K5, 8, 12, 16 or 20) on H4 as targets of MOZ acetylation. Further, to identify which histone H4 lysine residue is acetylated by MOZ, peptide binding assays were conducted with mono-acetylated histone H4 peptides to see whether blocking a specific residue would prevent acetylation of the peptide. Rather than identifying a specific target of acetylation, it appears that none of the mono-acetylated H4 peptides could be additionally acetylated by MOZ DPF+MYST (Figure 4.7). This suggests that H4Kac serves as a prevention signal for MOZ DPF+MYST binding to chromatin and the further acetylation of the same histone H4 tail. Whether this is because of the inability of H4Kac to interact with MOZ DPF remains to be established.



**Figure 4.7. Acetylation of histone H4 blocks further acetylation by MOZ.** *In vitro* HAT assays using 1.5  $\mu\text{g}$  HAT enzyme, 0.75  $\mu\text{g}$ /12.5  $\mu\text{M}$  histone peptides as substrate and [ $^3\text{H}$ ]-Acetyl-CoA as an acetyl group donor. HAT reactions were followed by filter binding and liquid scintillation counting. Error bars represent the average of a minimum of two data sets.

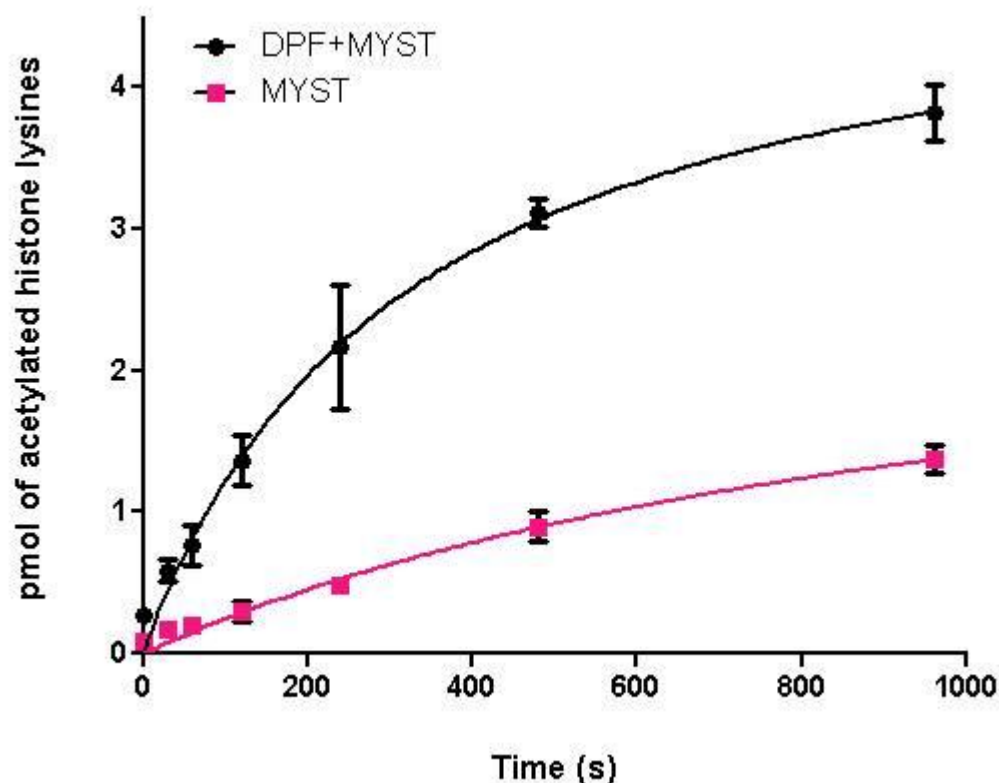
#### **4.4. Cooperative histone acetylation by the DPF+MYST module**

From previous acetylation assays using both peptides and core histones as substrates, it is evident that the acetylation activity of GST-MOZ DPF+MYST is higher than that of either domain alone (Figures 4.2, 4.3 & 4.5). These results may indicate the two domains cooperate to enhance the rate of histone acetylation. In order to test this hypothesis, time course HAT assays were conducted to compare the specific activity (amount of substrate the enzyme converts per mg of protein per unit of time) between core histones substrates, the [ $^3\text{H}$ ]-Acetyl-CoA acetyl group donor and GST-MOZ-DPF, -MYST, -DPF+MYST protein constructs. Protein quantities were normalised (Figure 4.1(B)) to ensure that the same amount of each construct was used in the assays. It is assumed that all three proteins are correctly folded due to significant binding and HAT activity in prior assays (Figure 3.11 & 4.2). In addition, the MOZ MYST domain has been structurally characterised by X-ray crystallography (Holbert et al., 2007) and the MOZ DPF domain has been structurally characterised by our laboratory in collaboration with Ingrid Dreveny (Chapter 3).

Filter binding assays were carried out as before, except that instead of taking a single point measurement after the reaction has reached steady-state, numerous samples were taken over a time course of 960 seconds in order to determine the rate of histone acetylation by each construct. To convert the scintillation counts (Dpm) into the more relevant measurement of pmol of acetylated histones, a fixed amount of tritium was spotted onto a filter and used as a standard.

The GST-MOZ MYST (Figure 4.8- pink) and GST-MOZ DPF+MYST (Figure 4.8- black) proteins are both active in this acetylation assay. Two main conclusions can be obtained from this kinetic plot; firstly, there is a clear difference in the extent of histone acetylation by the two constructs. For example at the fixed time value of 240 seconds, GST-MOZ MYST acetylates 0.48 pmol of histone substrate and GST-MOZ DPF+MYST acetylates 2.17 pmol of histone lysine residues; a 4.5 fold difference. Thus, a clear increase in the total acetylation output is observed in the presence of the DPF domain.

Secondly, the gradient of the plotted slopes differs significantly, where a steeper slope is observed for GST-MOZ DPF+MYST compared to GST-MOZ MYST (Figure 4.8). The slope represents the rate of acetylation catalysis, thereby depicting that the rate of histone acetylation by DPF+MYST is significantly faster than that of the MYST domain alone (Figure 4.8). In addition, these data show a difference in the rate at which the steady state is achieved by the two proteins. A linear relationship between time and activity can be observed for the GST-MOZ MYST protein plot, showing that the reaction has not yet begun to plateau and thus, has not reached a steady-state (Figure 4.8- pink plot). In comparison, the plot for GST-MOZ DPF+MYST can be divided into two sections. The plot is linear over the initial 240 seconds but begins to plateau between 240 and 960 seconds, indicating that the reaction has reached steady-state (Figure 4.8- black plot).



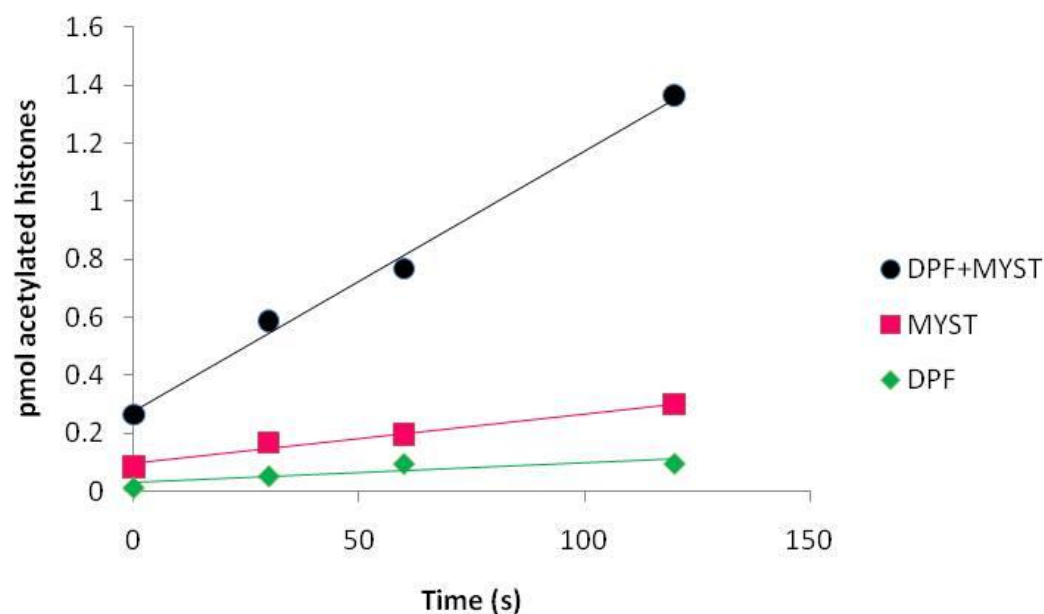
**Figure 4.8. The DPF domain enhances the HAT activity of MOZ.** *In vitro* HAT assays using 1.5  $\mu$ g of HAT enzyme, 10  $\mu$ g core histones as a substrate and [ $^3$ H]-acetyl-CoA as an acetyl group donor. Numerous samples were taken over a time course of 960s at intervals of 0, 30, 60, 120, 240, 480 and 960s. HAT reactions were stopped and samples were subjected to filter binding and liquid scintillation counting. Error bars represent the average of a minimum of two biological replicates.



To determine initial reaction rates for the DPF, MYST and DPF+MYST domain proteins, the first four data points for each construct were plotted. The slopes represent the initial rate of catalysis performed by the constructs and these data show that both the DPF and MYST domains alone undergo a much slower rate of acetylation compared to when both domains are present together (Figure 4.9(A)). Linear regression analysis was used to obtain the initial rate of catalysis, in pmol acetylated histones per second (pmol/s), from the slope of the fitted line. Subsequently, the rates were converted as ratios relative to the DPF+MYST rate of catalysis (Figure 4.9(B)). From this data it is evident that the DPF+MYST domain shows a 12.5- or 5-fold increase in specific activity over the individual DPF and MYST domains respectively.

These results could be attributed to the supplementary intrinsic HAT activity of the DPF domain, where both the DPF and MYST domains acetylate histones, enhancing the total output and rate of acetylation. Thus, the DPF+MYST module together is more likely to constitute the functional acetyltransferase domain of MOZ. In addition, the DPF could facilitate and augment catalysis by the MYST domain through its binding affinity for the substrate ( $K_d$ -dissociation constant), whereby it binds and docks onto the histone substrates that require acetylation. This would permit a specific targeted acetylation of the substrate rather than a random process, increasing the probability of successful catalysis.

(A)



(B)

Protein	Initial Rate of Reaction (pmol acetylated histones s <sup>-1</sup> )	Specific Activity (pmol acetylated histones s <sup>-1</sup> mg <sup>-1</sup> )	Ratio
DPF+MYST	0.0090	6000	1.00
MYST	0.0017	1133	0.19
DPF	0.0007	467	0.08

**Figure 4.9. The DPF domain enhances the rate of histone acetylation by MOZ.** *In vitro* HAT assays using 1.5 µg HAT enzyme, 10 µg core histones as a substrate and [<sup>3</sup>H]-Acetyl-CoA as an acetyl group donor. Samples at early time intervals 0, 30, 60, 120s were taken and subjected to filter binding and liquid scintillation counting. **(A)** Graphical representation of data. Slopes represent the initial rate of reaction of each MOZ construct. **(B)** Linear regression analysis was used to calculate the initial rate of reaction, in pmol acetylated histones s<sup>-1</sup>, from the slope of the fitted line. Enzyme specific activity values were calculated considering the 1.5µg of protein used in each assay and consequently, single DPF and MYST domain activities were converted into initial rate of acetylation fold ratios relative to DPF+MYST.

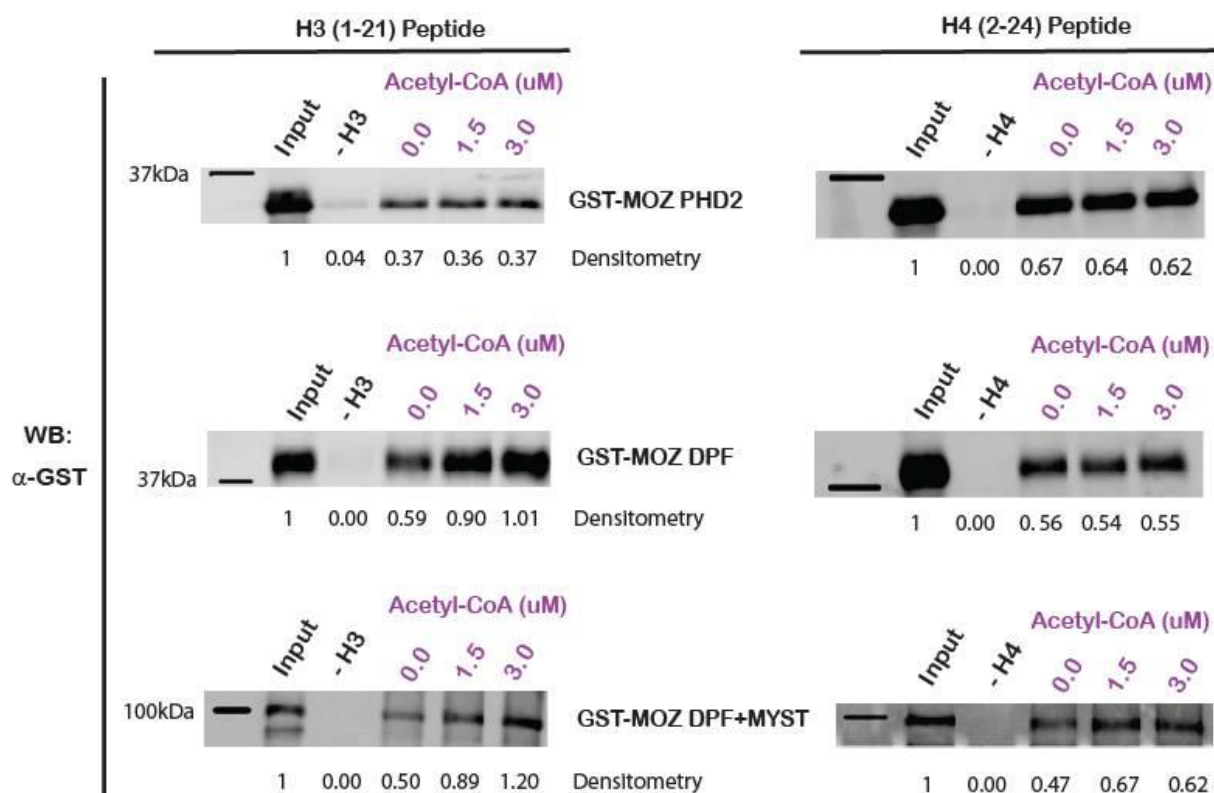
#### **4.6. Acetylation of histone H3 by MOZ DPF promotes its interaction with H3**

Binding studies revealed that acetylation of an unmodified histone H3 peptide stimulated interaction with MOZ DPF, whereas the hyperacetylation of histone H4 was not tolerated (Figure 3.11(B)). MOZ DPF and MYST domains also exhibit the capacity to acetylate both the H3 and H3/H4 substrates respectively (Figure 4.2(A)). We therefore investigated whether MOZ acetylation of the histone N-terminal tails could influence its binding preference for the substrate.

Histone peptide binding assays were performed with GST-MOZ DPF, -PHD2 or -DPF+MYST proteins and unmodified H3/H4 peptides in the absence or presence of acetyl-CoA. Western blots were used to detect the amount of GST protein retained on biotinylated peptide beads and assessed quantitatively using densitometry. The increasing presence of acetyl-CoA enhanced the binding of DPF and DPF+MYST to histone H3 in a dose-dependent manner (Figure 4.10- compare lanes 3 with 4 & 5, top and bottom left panel). No effect on the binding of PHD2 to histone N-terminal tails was observed (Figure 4.10- middle panel), consistent with its lack of acetyltransferase activity (Figure 4.3). These results indicate that acetylation of histone H3 by DPF or DPF+MYST promotes the interaction of MOZ to the H3 N-terminus. Thus, MOZ appears likely to promote its own recruitment to chromatin via the acetylation of and binding to histone H3.

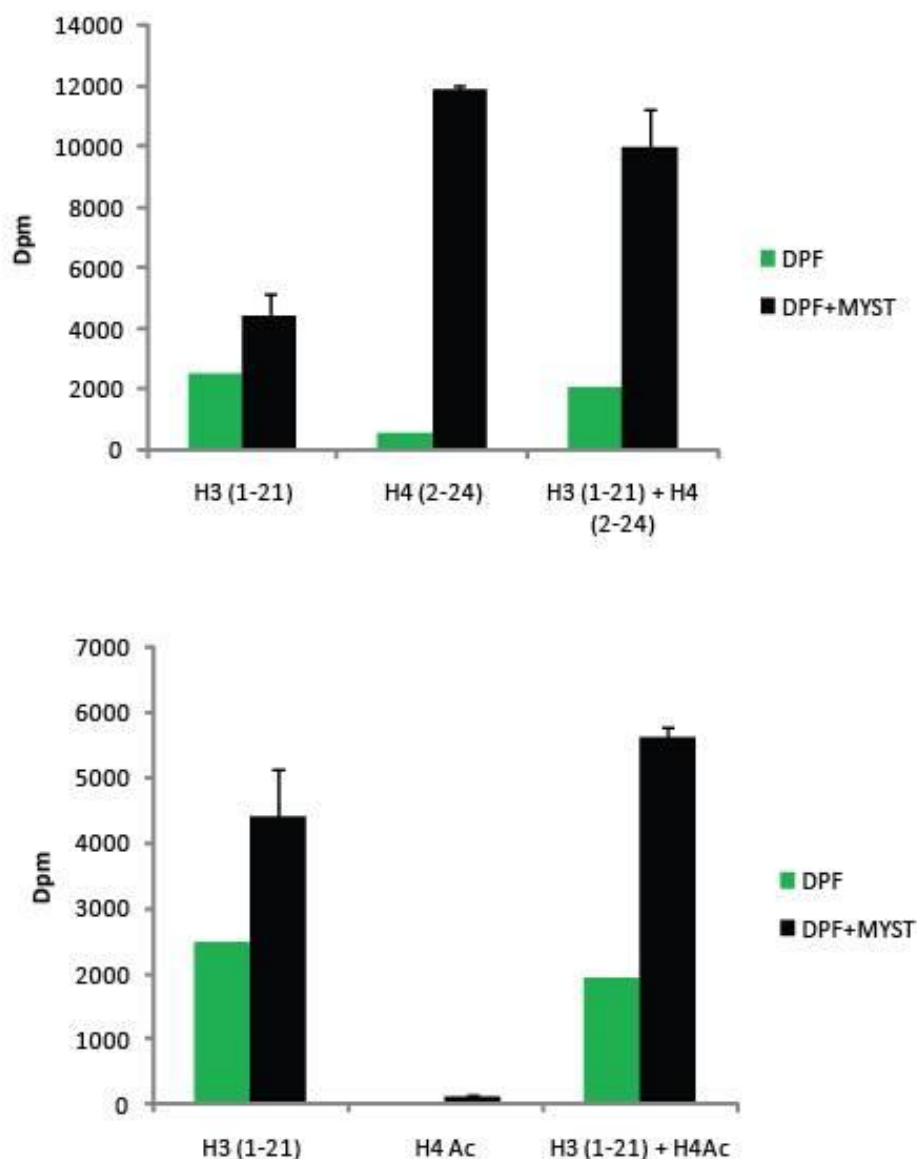
In addition, binding of the DPF and PHD2 proteins to histone H4 was unaltered by the presence of excess acetyl-CoA (Figure 4.10- compare lane 8 with 9 & 10, top and middle right panel). This result corroborates previous findings where

PHD2 lacks total histone acetyltransferase activity and MOZ DPF exhibits HAT activity towards histone H3 only and, is therefore, unable to acetylate histone H4 (Figure 4.3). In contrast, DPF+MYST, which can bind and acetylate unmodified H4, showed a net increase in binding in the presence of acetyl-CoA. Although DPF-MYST can acetylate H4, it is unable to form a stable association with acetyl-H4 as previously described. The observed increase may reflect an interim stabilisation through docking with the unmodified H4 peptide. This suggests that instead of functioning as an exit signal, rather pre-written or deposited H4 lysine acetylation serves to prevent the initial docking of MOZ to chromatin rich in this modification. Alternatively, acetyl-CoA may change the conformation of the DPF+MYST protein structure, rendering acetyl-H4 binding as favourable.



**Figure 4.10. Acetylation of histone H3 by MOZ DPF promotes its interaction with H3.** *In vitro* HAT assays were conducted with 1.5  $\mu$ g of the indicated GST-MOZ fusion protein, 0.75  $\mu$ g/12.5  $\mu$ M biotinylated unmodified histone H3 or H4 peptides as substrate and increasing concentrations of cold Acetyl-CoA (0, 1.5 and 3  $\mu$ M) as an acetyl group donor. Subsequently, histone peptide pulldowns were carried out where samples were incubated overnight in binding buffer and subjected to immunodetection with  $\alpha$ -GST. Densitometry readings were calculated, relative to the protein input, using the Image J software.

Hyperacetylation of histone H4 influences the ability of the DPF domain to bind other H3 PTMs, but it is unclear whether this dominant inhibitory affect is also extended to its HAT activity. In order to address this question, single peptides or an equimolar mixture of peptides were used as substrates in an *in vitro* HAT assay with GST-MOZ DPF and DPF+MYST. Firstly, as a control assay, peptides representing the N-terminal tails of both unmodified H3 and H4 were incubated together in the HAT assay. As shown in Figure 4.11, the presence of the H4 unmodified peptide does not alter the ability of the DPF to acetylate histone H3 (Figure 4.11(A)- compare green bar 1 with 3). In addition, the total acetylation level with DPF+MYST and a 50:50 ratio of H3 and H4 peptides appears not to be affected (Figure 4.11(A)- compare black bars 1 & 2 with 3). However, as both peptides are targets for the DPF+MYST HAT activity, it is not clear whether the acetylation level of H3 and H4 individually is affected in the presence of the other peptide. Secondly, the unmodified histone H3 peptide was incubated with the hyperacetylated histone H4 peptide. As only the H3 peptide is a target for acetylation in this assay, due to the use of a pre-acetylated H4 peptide (Figure 4.11(B)- compare green & black bars 1 & 2), a direct comparison can be made between the level of H3 acetylation in the absence and presence of the H4Kac peptide. As shown in Figure 4.11, the level of H3 acetylation for both the DPF and DPF+MYST proteins is unchanged in the presence of the H4Kac peptide (Figure 4.11(B)- compare green & black bars 1 & 3). This suggests that pre-acetylated histone H4 exerts a dominant inhibitory affect over binding of MOZ DPF to chromatin, but does not influence its ability to acetylate neighbouring H3 N-terminal tails.



**Figure 4.11. H4 acetylation does not affect the acetylation of histone H3.** *In vitro* HAT assays using 1.5  $\mu$ g HAT enzyme, a total mass/concentration of 0.75  $\mu$ g/12.5  $\mu$ M histone peptides (**A**) H3 (1-21) and H4 or (**B**) H3 (1-21) and pre-acetylated H4Kac as substrate(s) and [ $^3$ H]-Acetyl-CoA as an acetyl group donor. HAT samples were then subjected to filter binding and liquid scintillation counting.

### **4.7. The MOZ DPF domain is required for AML-1 mediated transcription**

MOZ functions as a transcriptional coactivator for several transcription factors, particularly factors involved in the regulation of haematopoiesis. MOZ has been shown to interact with AML1 and stimulate the transcription of AML-dependent genes (Kitabayashi et al., 2001a). Intriguingly, it has been shown that mutations that disable the MYST HAT domain did not affect the coactivator properties of MOZ, but rather the function of a C-terminal transactivation domain (serine-, proline & glutamine-, methionine-rich regions) and the N-terminal H1/5 domain was required for transcriptional activity (Kitabayashi et al., 2001a). To investigate whether the MOZ DPF domain was required for the stimulation of AML1-mediated transcription, luciferase reporter assays were conducted. U2OS cells were transiently transfected with the AML1-responsive luciferase reporter plasmid (pT109-3x AML1-luc), the AML1 and FLAG-MOZ coding expression vectors and the  $\beta$ -galactosidase internal control plasmid. Luciferase activity was assayed 48 h post transfection using the Dual-light Luciferase and  $\beta$ -Galactosidase reporter gene assay system (Applied Biosystems) and Luminometer Orion (Berthold).

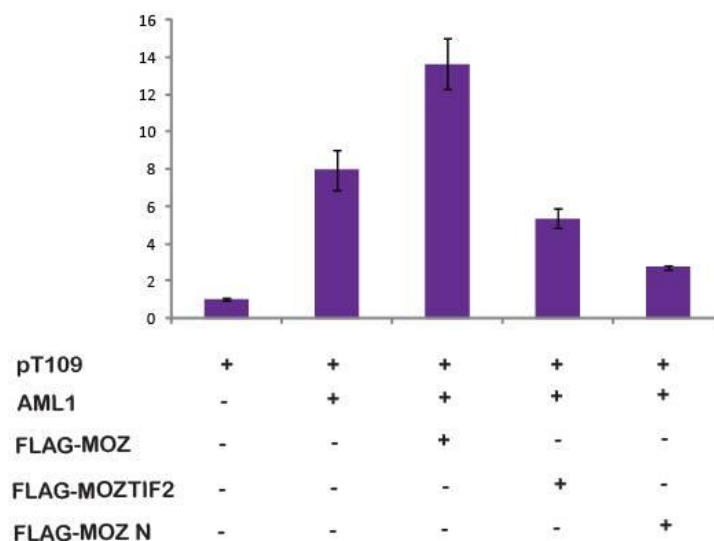
AML1 was cotransfected with either the full-length MOZ construct, MOZ N (breakpoint fusion protein, lacking the C-terminus), MOZ-TIF2 oncogenic fusion protein or empty vector and were all tested in the luciferase reporter assay to examine their affect on AML1-mediated transcription. Full-length MOZ was able to activate transcription at the AML1 promoter (Figure 4.12(A)- compare bars 2 & 3). These findings are consistent with previously published results from our group (Collins et



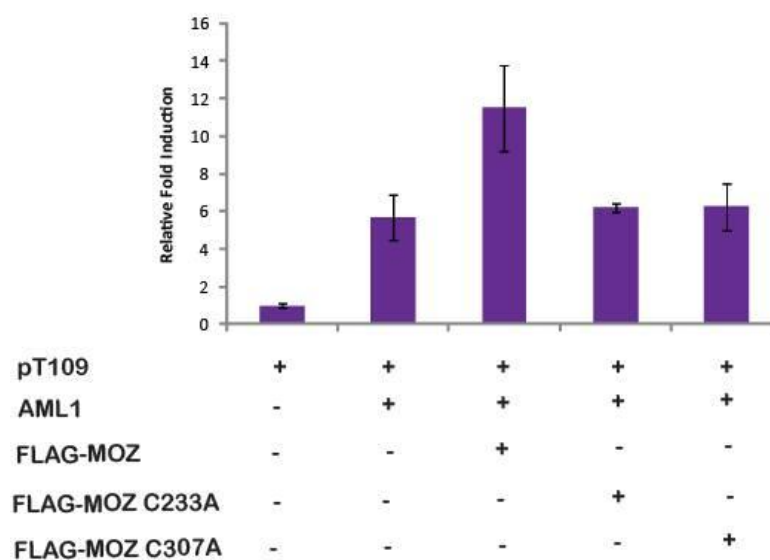
al., 2006), underscoring the necessity of the C-terminal domains for transcriptional activation. Single DPF mutants were therefore created in the full-length MOZ expression plasmid for testing in subsequent reporter assays. Both the PHD1 (C233A) and PHD2 (C307A) mutants abolished the transactivation potential of MOZ at the AML1 promoter (Figure 4.12(B)- compare lane 3 with 4 & 5), suggesting that the DPF domain is required for coactivator function. These DPF mutants should not disrupt activation through altering the direct interaction of MOZ with AML1, as the two binding regions for AML1 have been mapped outside of this domain (Kitabayashi et al., 2001a). Mutating zinc-coordinating cysteine residues is likely to disrupt the structure of the DPF, however little is known about the affect these mutants may have on the overall protein structure. Yet, immunofluorescence data accentuates that the protein mutants are expressed in these cells (Figure 4.13). These factors imply that the abrogation of AML1-mediated transcription is caused by the disruption of the MOZ DPF domain, albeit via a mechanism that is independent of the AML1-MOZ interaction. The PHD domain has also been cited as functioning as a protein-protein interaction domain (Bienz 2006) and possibly mutating the MOZ DPF prevents the recruitment of additional factors required for the stimulation of AML1-mediated transcription. Alternatively, it has been shown that the role of MOZ HAT activity in gene regulation is dependent on promoter contexts, where the MYST domain is required for NF- $\kappa$ B-dependent transcription (Chan et al., 2007), but not for AML1-dependent transcription (Kitabayashi et al., 2001a). This suggests that H4 acetylation by MOZ MYST is not required for coactivator function but the DPF domain is. Earlier results demonstrate that the MOZ DPF can function as a chromatin binding module (Figure 3.11) or as a mild H3-specific acetyltransferase (Figure 4.2).

Thus, its HAT activity may also be required for coactivator properties at a specific subset of promoters. The abolished transactivation of AML1-mediated transcription could therefore, be a result of a lack of DPF-mediated histone H3 acetylation or a result of the loss of ability to bind chromatin.

(A)



(B)



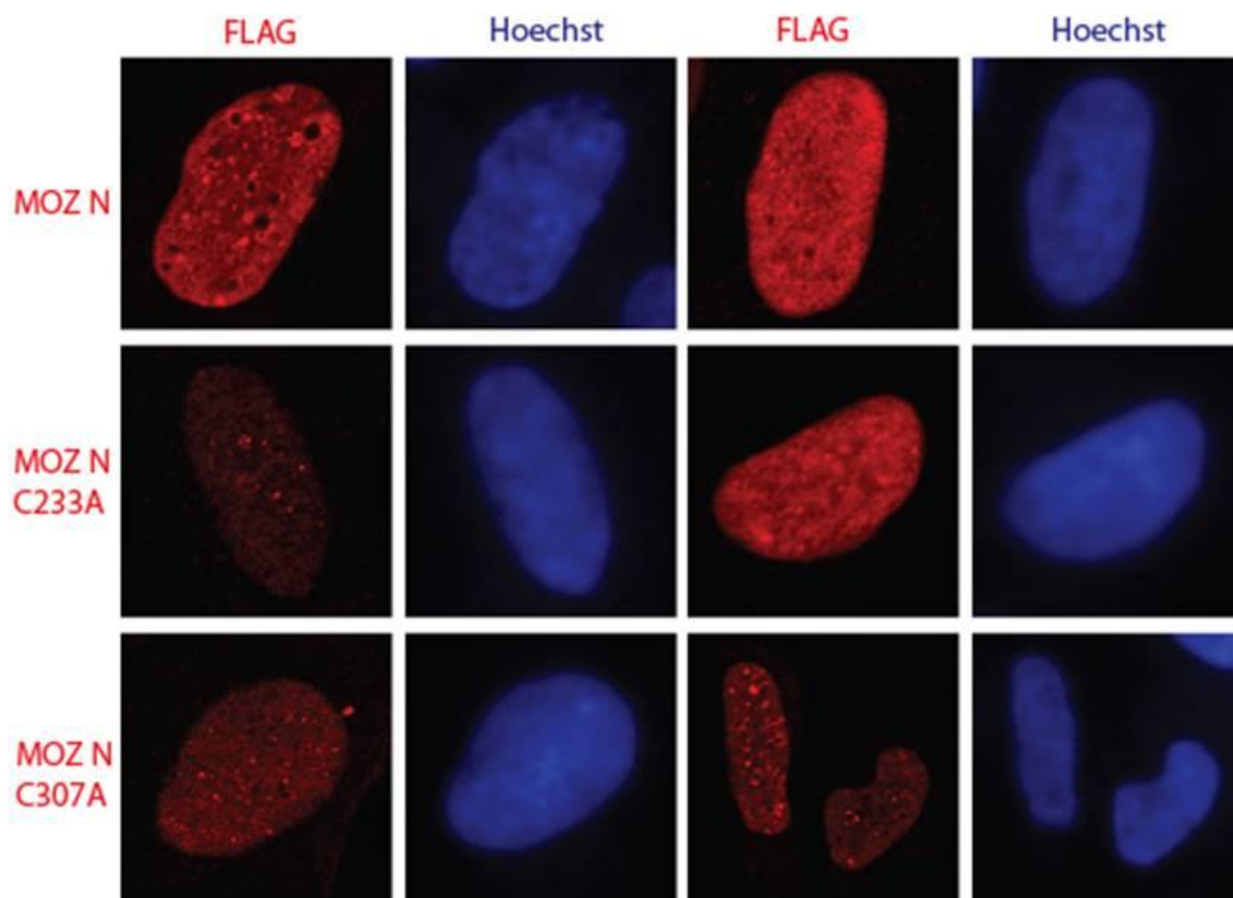
**Figure 4.12. MOZ DPF mutants inhibit transcription activation of the AML1 reporter.** U20S cells were transiently transfected with 500 ng of pT109-3x AML1-luc, 100 ng of AML1 and  $\beta$ -GAL and 500 ng FLAG-MOZ construct in 12-well plates. Luciferase activity was assayed as described in Materials and Methods for (A) MOZ wild-type or (B) MOZ DPF mutant constructs. Reporter activation is represented as a fold induction over the control value (reporters in the absence of coactivators). The data are presented as averages and standard errors of the means from triplicate samples.

### 4.8. Mutation of MOZ DPF does not affect its subcellular localisation

The mutation of the DPF domain within the context of FLAG-MOZ led to a loss of coactivator function in reporter assays. To determine if this resulted in the destruction or mislocalisation of MOZ, we examined the subcellular localisation of MOZ using IF staining.

Previous reports from our group showed that MOZ was concentrated into subnuclear foci or speckles and was excluded from nucleoli. Moreover, co-staining for other nuclear proteins divulged that MOZ foci were distinct from splicing speckles, Cajal bodies and PML bodies, suggesting that MOZ inhabits a separate nuclear subdomain (Kindle et al., 2005). FLAG-tagged MOZ N and MOZ N DPF mutants were transiently transfected and over-expressed in HEK-293 cells. Cells were then stained with anti-FLAG and an Alexa-594-conjugated secondary antibody. Cell nuclei were also stained with the Hoechst 33258 nuclear stain.

The MOZ N WT, MOZ C233A and MOZ C307A proteins showed similar distribution patterns within the nucleus of transfected HEK293 cells. All three proteins exhibited punctate nuclear staining where proteins are concentrated into discrete foci (Figure 4.13). This suggests that although the DPF is essential for the engagement of chromatin with regards to histone PTMs, other domains within MOZ may provide additional contacts with chromatin or other nuclear structures.



**Figure 4.13. Subcellular localisation of exogenously expressed MOZ proteins.** HEK-293 cells were seeded onto coverslips and transfected after 24 hours with expression plasmids encoding WT FLAG-MOZ N, C233A or C307A FLAG-MOZ N mutants. After 48 hours post-transfection, cells were fixed in 4 % paraformaldehyde, permeabilised and stained with mouse anti-FLAG (red). Staining was followed by incubation in an anti-mouse Alexa594 secondary antibody. Hoechst 33258 (blue) was used as a stain of DNA.

### 4.9. Summary

In Chapter 3 it was shown that the MOZ DPF functions as a chromatin binding module by engaging specifically with the N-terminal tails of histones H3 and H4. MOZ appears to be incapable of forming a stable interaction with the trimethylated H3K4 and hyperacetylated form of histone H4. The inability to bind acetylated forms of H4 is intriguing, given that the MOZ MYST domain has been reported to target histone H4 acetylation *in vitro* (Champagne et al., 2001) (Fraga et al., 2005). However, isolated ING5/MOZ complexes, which also contain the proteins BRPF1 and EAF6, have been reported to target H3 acetylation in particular H3K14. It was suggested that the substrate specificity of MOZ is altered by ING proteins (Doyon et al., 2006). More recently, Voss *et al.* reported that MOZ ablation leads to hypoacetylation at H3K9, whereas neither H3K14 nor H4 acetylation was greatly affected (Voss et al., 2009). This presented a conundrum regarding MOZ histone substrate specificity, which appears to be different in opposing studies. Previous data from our lab using MOZ MYST domain (510-810) in HAT assays agreed with the finding that H4 is the major target of this domain, at least *in vitro*. It was therefore decided to assess whether the presence of the DPF would influence the substrate targeting of the MOZ MYST domain. To investigate this, we performed *in vitro* HAT assays using DPF, MYST or DPF+MYST constructs. Interestingly, these preliminary experiments uncovered a function for the MOZ DPF as a mild H3-specific acetyltransferase. Moreover, assays conducted with the individual PHD finger proteins (PHD1 and PHD2) elucidated that this acetyltransferase activity was an intrinsic property of the composite DPF module. To our knowledge, this constitutes

the first description of a possible acetyltransferase function attributed to a DPF domain or PHD finger. To progress with this study, it will be important to determine whether MORF or the DPF1/2/3 proteins also share the ability to acetylate histones and whether they have similar or distinct substrate preferences.

While PHD motifs are well established as independent protein ligand interaction domains (Bienz, 2006), a recent study has reported an enzymatic activity associated with the PHD finger, e.g. as a functional E3 ligase for SUMO molecules. The PHD domain of the KAP1 corepressor protein functions as a unique E3 ligase that directs the intramolecular SUMOylation of the adjacent bromodomain; an event required for KAP1-mediated repression of target genes (Ivanov et al., 2007). Thus, the PHD domain appears to act as a mediator module to stimulate and stabilise the functional outputs of the protein.

As we have shown that MOZ DPF has dual functionality, as a mild histone H3-specific acetyltransferase and a chromatin binding module, we investigated how this would impact on the activity of the MYST domain. We therefore conducted *in vitro* HAT assays in order to explore whether the DPF affects the acetyltransferase function of MOZ. With regards to substrate specificity, the MOZ MYST domain was shown to preferentially acetylate histone H4 as previously reported. However, a construct containing both the DPF and MYST domains appears to have an extended range of histone substrates, targeting H3, H4 and possibly H2A. Thus, the full-length MOZ protein contains two distinct acetyltransferase activities; the DPF which appears to selectively acetylate H3 and the MYST for which H4 is the main target. However, these activities may be modulated by other proteins *in vivo*, such as by those proteins found in complex with MOZ including ING5 and BRPF1 (Doyon et

al., 2006). Substrate specificity may also change depending on the form that the histone substrate is presented. For example, the MORF MYST domain was shown to acetylate free histones H3 and H4, but only has activity toward nucleosomal histone H3 (Ullah et al., 2008).

In addition to influencing the substrate specificity of MOZ, the presence of the DPF domain was also found to enhance the kinetics of histone acetylation by MOZ. Kinetic studies revealed that the DPF+MYST construct had significantly increased total acetylation output, as well as performing at a considerably faster rate of catalysis compared to either module functioning separately. Thus, it is likely that the DPF+MYST domains together constitute the MOZ functional acetyltransferase domain. Together they exhibit a 12.5- or 5-fold increase in specific activity compared to the individual DPF and MYST counterparts respectively. At present the DPF exhibits two functions and which of these functions, if not both, is responsible for the observed enhancement of HAT activity has not yet been resolved. Possibly, the combined acetyltransferase activities of the DPF+MYST protein towards an extended range of histone substrates could lead to the augmentation of the rate of catalysis. Alternatively, the DPF could strengthen the binding affinity of the MOZ protein for the substrate, thus facilitating the docking onto the histone that requires acetylation. To determine this, fluorescence polarisation or isothermal titration calorimetry experiments could be carried out to measure the  $K_d$  (dissociation constant) value. The  $K_d$  is a measurement of the strength of binding between receptor and ligand, thus this would elucidate whether the DPF+MYST bound with a greater affinity to the histone substrate than the MYST domain alone. Other examples of interplay between histone ‘reader’ and ‘writer’ modules have been documented in the literature. One such



example is the synergy between the bromodomain and HAT domain of the acetyltransferase p300. Studies by Chen *et al.* reveal the importance of the bromodomain for p300 to exert its HAT activity, whereby the bromodomain stimulates p300 HAT activity (Chen et al., 2010) by recognising acetylated lysine residues and stabilising the association of p300 to specific chromatin sites (Zeng et al., 2008b).

As the DPF exhibits dual function; as an acetyltransferase and a H3/H4 PTM sensor, we next investigated how existing PTMs might affect MOZ HAT activity. We performed *in vitro* HAT assays using a range of peptides, which included both modifications that promoted and those that disrupted DPF binding to histones, to test whether MOZ DPF binding to histones would influence the ability of MOZ to acetylate H3 and H4 N-terminal tails. We found that the MOZ DPF was able to acetylate the N-terminus of histone H3 (1-21), but not a histone H3 internal sequence (21-44). This implies that lysine residues within the first 21 amino acids of the histone H3 tail are targets for MOZ DPF acetylation. Trimethylation at K4 and K9 prevented the acetylation of the H3 peptide by the MOZ DPF. We have previously demonstrated that the MOZ DPF cannot bind H3K4me3 and that it is excluded from chromatin enriched with this modification, hence it may not be able to acetylate this peptide. The MOZ DPF is able to bind the H3K9me3 modification and so this indicates K9 as a likely target of MOZ DPF acetylation. Indeed, immunodetection analysis of HAT assays supported H3K9 as a preferred substrate for MOZ DPF acetylation.

The MOZ MYST domain acetylated histone H4 to a greater extent than H3 (1-21) and the combined DPF+MYST construct showed strong acetylation of both

H3 and H4. For both constructs, H3 acetylation was partially reduced by H3K9me3 but completely blocked by trimethylation at H3K4. A partial reduction of H3 acetylation in the presence of K9me3 suggests that, as well as K9 being an acetylation target, the MOZ MYST and DPF+MYST constructs are able to acetylate other lysine residues within the H3 peptide. In accordance, immunodetection analysis of HAT assays supported both H3K9 and H3K14 as preferred substrates for MOZ DPF+MYST acetylation.

All three constructs, MOZ-DPF, -MYST and -DPF+MYST were unable to acetylate a histone H3 peptide trimethylated at K4. The diminished activity for this particular substrate could in part be due to the intolerance of MOZ DPF for H3K4me3 binding. This suggests that MOZ is unable to tether and stabilise the association of the protein onto peptides containing H3K4me3, preventing further acetylation of lysine residues on the same histone tail. To investigate these findings in an *in vivo* setting, it would be desirable to conduct the ChIP-Seq protocol in order to identify genome-wide MOZ binding sites and their association with PTMs. MOZ-tagged genomic sites could then be used to map the proximity of complexes to promoters and further, be correlated with one or more histone PTMs. The MYST domain was also unable to acetylate the H3K4me3 peptide. This suggests that modification of H3K4 could possibly prevent the entry of the substrate into the active site, either by size or charge exclusion or a change to the conformation of the histone peptide, impeding acetylation of the histone lysine substrate.

We have established that the MOZ DPF and DPF+MYST domains acetylate the H3 and H3/H4 substrates respectively. We have also previously shown through binding studies that acetylation of H3 stimulated the interaction with MOZ DPF,

whereas conversely the acetylation of H4 prevented the interaction with MOZ DPF. Therefore, we next investigated whether MOZ-mediated acetylation of H3 and H4 N-terminal tails could influence its binding preference for the substrate, and thus regulate its own recruitment or exit from chromatin. We performed histone peptide binding assays with H3 and H4 peptides in the absence or presence of acetyl-CoA, and found that the increasing presence of acetyl-CoA enhanced the binding of the DPF and DPF+MYST to histone H3. This indicates that MOZ promotes its own recruitment to chromatin via the acetylation and binding of histone H3. In addition, DPF+MYST, which can bind and acetylate H4, also showed an increase in binding in the presence of acetyl-CoA. This increase is likely to represent binding to the interim unmodified H4 peptide, suggesting that MOZ-mediated histone H4 acetylation does not function as a release mechanism for MOZ from the H4 peptide. In contrast, peptide acetylation assays show that mono-acetylated peptides could not be additionally acetylated by MOZ DPF+MYST. This implies that pre-written H4Kac serves as a prevention signal for MOZ to avert binding to chromatin and further acetylation of the same H4 tail. However, this H4Kac signal was not shown to influence the ability of MOZ to acetylate neighboring H3 N-terminal tails. All our studies, thus far, investigate the affect of existing PTMs and the DPF on MOZ HAT activity in the context of immobilised peptides and free histones. The next stage would be to extend the study to a more physiologically relevant substrate such as mono- and oligo-nucleosomes.

The p300 bromodomain is not only required for stimulating p300 HAT activity, but is also important for p300 to function at full coactivator activity in the activation of p300-dependent genes (Chen et al., 2010). This led to the investigation

as to whether the DPF module could also affect other functional properties of the MOZ protein. MOZ functions as a coactivator for several transcription factors, especially those with hematopoietic specificity. MOZ has been shown to be essential for AML1-mediated transcription, whereby the C-terminal transactivation domain and the H1/5 domain were required for transcription activity of AML1-dependent genes. Interestingly, it was shown that the MYST domain was not important for coactivator function (Kitabayashi et al., 2001a), implying that H4 acetylation is not required. We carried out luciferase reporter assays, using the AML1-responsive luciferase reporter, to determine whether the DPF domain was important for the stimulation of AML1-mediated transcription. Both the PHD1 and PHD2 mutants diminished the transactivation potential of MOZ at the AML1 promoter, suggesting that the composite DPF domain is required for coactivator function. It is possible that the abolished transactivation of AML1-mediated transcription could be as a result of the loss of the ability to bind chromatin or the lack of DPF-mediated histone H3 acetyltransferase activity. Further, to determine whether the mutation of MOZ resulted in the mislocalisation of the protein, we performed immunofluorescence studies. Mutants revealed similar distribution patterns to the wild-type protein, thus it appears that other domains within MOZ may provide additional contacts with chromatin or other nuclear structures. Indeed, the H1/5 domain has been shown to be important for the nuclear localisation of MOZ (Kitabayashi et al., 2001a).

In summary, we have uncovered a second function for the MOZ DPF in histone acetylation. Preliminary results indicate that the DPF domain acts as a H3-specific acetyltransferase, which acts to extend the range of histone substrates and enhance the acetyltransferase kinetics of MOZ. Results also begin to explore the

affect of DPF PTM binding preference on the HAT activity of MOZ, however it is clear that more experimental data is required to fully elucidate the mechanisms of how the DPF regulates MYST activity.

## **Chapter Five: Results**

### **MOZ interacts with other chromatin modulators**

**5.1. Introduction**

In the previous chapters, it was shown that the MOZ DPF domain can associate with chromatin enriched in the tri-methylated H3K9 PTM, but not H3K4me3, as supported by peptide binding studies (Figure 3.11). H3K4me3 is generally considered a marker of active genes, whereas H3K9me3 is a hallmark of heterochromatin, a densely compacted chromatin conformation associated with transcriptional silencing. The SUV39H1 lysine methyltransferase enzyme specifically mediates the trimethylation of histone H3 lysine 9, which consequently creates a binding site for the protein machinery involved in promoting the formation of heterochromatin (Melcher et al., 2000). The heterochromatin HP1 proteins can bind H3K9me3 enriched chromatin and additionally interact with SUV39H1. Thus, in general, SUV39H1 methylation of histone H3K9 creates a high-affinity binding site for the HP1 chromodomain, in addition to direct recruitment of HP1 via the chromoshadow domain. Thereby, these proteins function in at least a three component complex to facilitate the spreading of higher order chromatin states (Fanti and Pimpinelli, 2008).

As described in chapter 4, MOZ interacts with AML1 (Kitabayashi et al., 2001a) and cooperatively stimulates the expression of AML1-dependent target genes (Collins et al., 2006). AML1 also forms a complex with SUV39H1, which conversely abrogates the DNA binding and transactivating properties of AML1 (Chakraborty et al., 2003). The association between AML1 and SUV39H1 provides a molecular mechanism for gene repression mediated by AML1 and further, the distinct

interactions with these coregulators determine the strength of AML1 in promoter regulation.

SUV39H1 activity is regulated by the acetylation of lysine 266 in its catalytic domain (Vaquero et al., 2007). The deacetylase SIRT1 interacts directly with and deacetylates SUV39H1, enhancing levels of SUV39H1 enzymatic activity and resulting in the increased trimethylation of histone H3 lysine 9 (Vaquero et al., 2007). The identity of the KAT enzyme responsible for the acetylation of SUV39H1 and the successive regulation and impairment of SUV39H1 activity remains elusive.

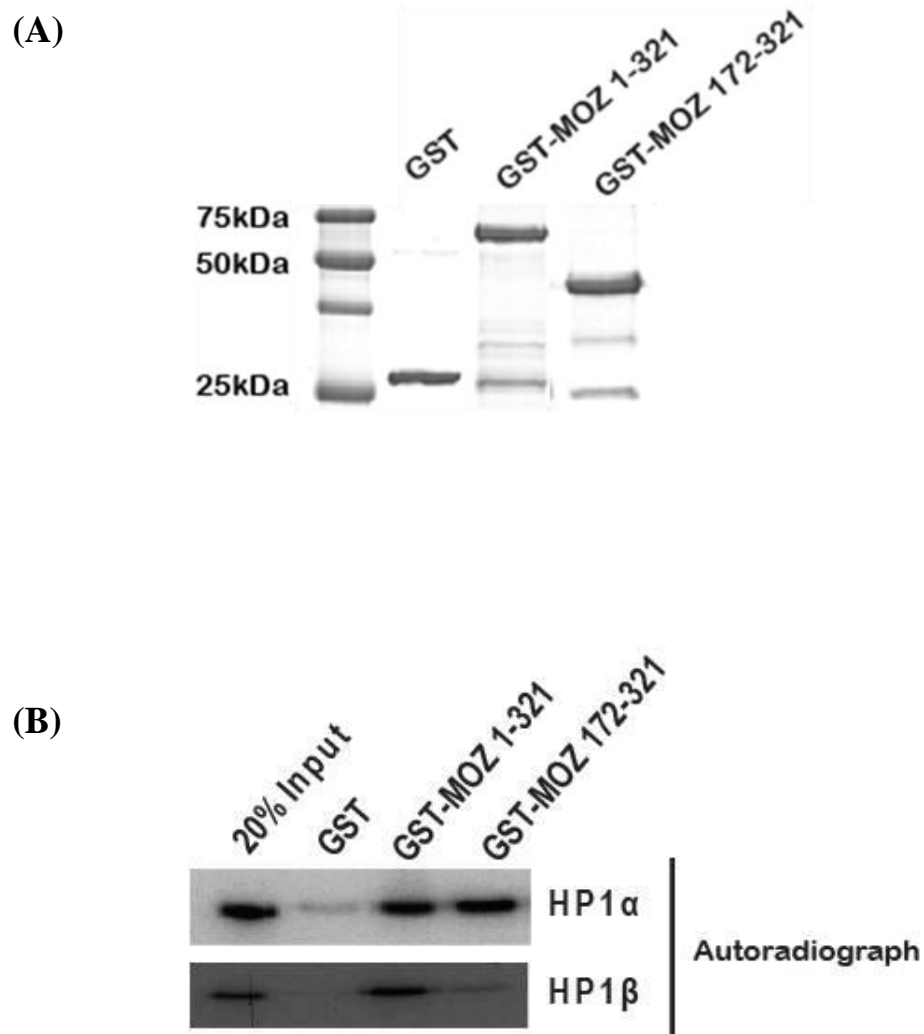
MOZ associates with the repressive H3K9me3 modification *in vitro*. The literature details that both MOZ and SUV39H1, the methyltransferase responsible for the catalytic deposition of this mark, function as coregulators of AML1-dependent gene expression. Moreover, the link between SUV39H1 and KAT enzymes is strengthened as the regulation of its own activity is governed by post-translational acetylation. This chapter therefore, describes work carried out to examine whether the MOZ KAT can associate with proteins involved in the formation of heterochromatin and the regulation of gene silencing, including HP1 and SUV39H1. Moreover, preliminary studies begin to uncover the functional ramifications of such interactions.

### 5.2. MOZ interacts with HP1

In this study evidence has been provided that MOZ associates with H3K9me3. HP1 is a well-established H3K9me3 binding protein, thus we first investigated whether MOZ might interact with HP1 proteins.



The GST-MOZ N-terminus (1-321) and DPF (172-321) fusion proteins were expressed and purified as described in previous sections (section 3.4 and 4.2). Protein concentrations were normalised by SDS-PAGE and Coomassie staining to ensure that equal amounts of GST proteins were used in pulldown assays (Figure 5.1(A)). The three HP1 isoforms, HP1 $\alpha$ ,  $\beta$  and  $\gamma$  (provided as gifts from Dr. B Le Douarin) were *in vitro* translated using  $^{35}\text{S}$ -methionine. While HP1  $\alpha$  and  $\beta$  isoforms were successfully produced to give proteins of the expected molecular weights (as shown in the input lanes-Figure 5.1), no band was obtained for the HP1 $\gamma$  isoform. Although the sequence indicated that the HP1 $\gamma$  cDNA insert was of the correct orientation and sequence, we were unable to detect *in vitro* translated protein. Thus, it was not possible to test the interaction of HP1 $\gamma$  and MOZ in this *in vitro* system. As shown in Figure 5.1, no interaction was observed between GST and HP1  $\alpha$  and  $\beta$  proteins, indicating that interaction does not occur through the GST tag (Figure 5.1(B)- lane 2). HP1 $\alpha$  was pulled down, to an equal degree, by both the MOZ N-terminus and the DPF domain (Figure 5.1(B)- top panel, lanes 3 & 4). The HP1 $\beta$  isoform also showed an interaction with the MOZ N-terminus, although a much weaker interaction with the DPF domain alone was observed (Figure 5.1(B)- bottom panel, lanes 3 & 4). This result suggests that HP1  $\alpha$  and  $\beta$  isoforms can interact directly with the N-terminus of MOZ *in vitro*, and that the HP1 $\alpha$  binding site is likely to be located adjacent to or within the DPF domain of MOZ (amino acids 172-321).

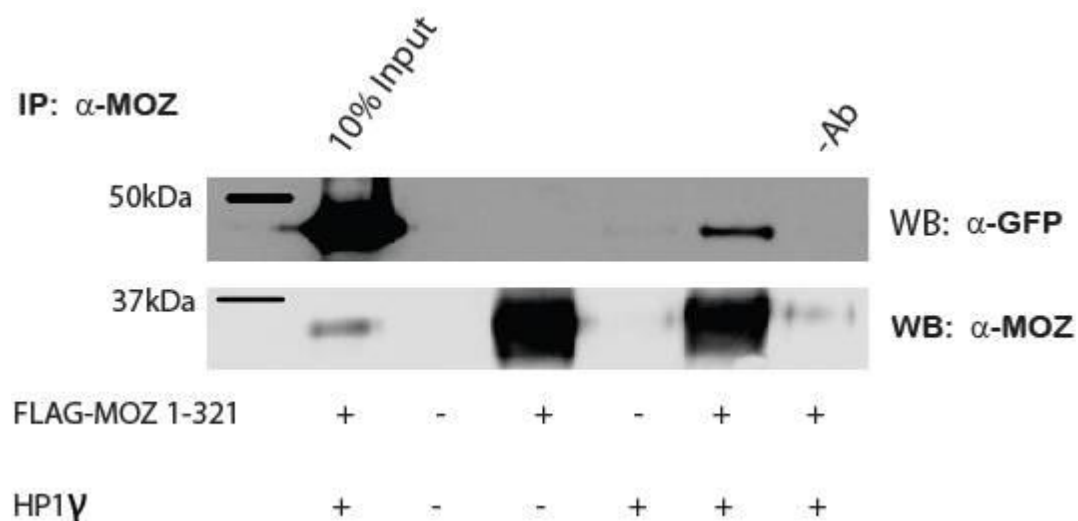


**Figure 5.1. HP1  $\alpha$  and  $\beta$  interaction with the N-terminus of MOZ *in vitro*.** (A) Normalisation of GST-MOZ constructs concentration by SDS-PAGE and Coomassie staining. (B) Pulldown assays with *in vitro*-translated  $^{35}\text{S}$ -methionine labelled HP1  $\alpha$ ,  $\beta$  and GST-MOZ 1-321 (H1/5 & DPF domains) and 172-321 (DPF domain) followed by autoradiograph analysis.

To validate that MOZ and HP1 can interact with one another in mammalian cells, co-immunoprecipitations were performed with the FLAG-MOZ N-terminus (1-321) and HP1 $\gamma$  proteins expressed in HEK293 cells. Attempts to co-immunoprecipitate HP1 $\alpha$  and HP1 $\beta$  using a commercial antibody, raised to detect all three isoforms, were unsuccessful. However, one of the constructs we obtained consisted of HP1 $\gamma$  fused to the YFP protein, so we were able to use this construct in co-IP and immunofluorescence studies.

FLAG-MOZ 1-321 and YFP-HP1 $\gamma$  exogenous proteins were transiently transfected into HEK293 cells and 48 hours post-transfection, cells were harvested and lysed to release soluble proteins. Input lanes verify the expression of each transfected protein in the cell lysates (Figure 5.2(A)- lane 1). FLAG-MOZ protein was then immunoprecipitated from cell lysates using the MOZ (127A) antibody and tested for HP1 $\gamma$  interaction by immunodetection with the GFP antibody (Figure 5.2(A)). To determine the efficiency of the IP, the blot was also probed with the MOZ (127A) antibody. As expected, the MOZ (127A) antibody immunoprecipitated a band at approximately 35kDa (Figure 5.2(A)- bottom panel) only in lanes that expressed FLAG-MOZ. A protein at ~45kDa was detected when both FLAG-MOZ 1-321 and YFP-HP1 $\gamma$  were co-expressed, but not in control mock transfected lysates or no antibody IPs (Figure 5.2(A)- top panel, compare lanes 2 & 6 with 5). The identity of the HP1 protein was confirmed using the GFP antibody. This result shows that the MOZ N-terminus can form a complex with HP1 $\gamma$  *in vivo*.

(A)



(B)

MOZ	184	FPSSL	PPVSL	LPHEK	198
MORF	191	FPSSL	PPVSL	LPHEK	205
HP1α	156	ANVKC	PQIVIAFYEE		170
HP1γ	146	ANMKC	PQIVIAFYEE		160
TIF1α	670	STHKV	PPVML	EPIRI	684
TIF1β	481	LLRKV	PRVSL	ERLDL	495
Su (var) 3-7	308	DAQDT	PPVMM	NGDVD	322
CAF-1 (p150)	215	FKGKV	PMVVL	QDILA	229

**Figure 5.2. MOZ interacts with HP1γ *in vivo*.** (A) HEK293 cells were mock transfected or transiently cotransfected with YFP-HP1γ and FLAG-MOZ 1-321. After 48 h post-transfection, cells were lysed and the proteins were coimmunoprecipitated with α-MOZ. Proteins were separated by electrophoresis and subjected to immunoblotting using α-MOZ and α-GFP antibodies. (B) Alignment of an N-terminal MOZ/MORF pentapeptide sequence with other HP1-associated protein sequences containing the pentapeptide consensus (pink). Numbers denote the positions of the amino acids shown.

Many HP1-binding proteins contain a consensus pentapeptide sequence, which is sufficient for specific interaction with the HP1 chromoshadow domain. This consensus comprises a proline at position 1, valine at position 3 and hydrophobic ( $\psi$ ) residues at positions 4 and 5 (Cowieson et al., 2000). As the MOZ N-terminus has been shown to interact with HP1, we next analysed the MOZ sequence (1-321) for this PxV $\psi\psi$  motif. As shown in Figure 5.2, the MOZ sequence proline 189 to leucine 193 fits this consensus (PPVSL), which is also fully conserved in the homolog MORF (Figure 5.2(B)). An alignment of MOZ and MORF sequence with PxVxL motifs in other HP1-associated proteins is shown in Figure 5.2(B). The GST pulldown data shown in this study indicated that the MOZ sequence 172-321 was sufficient to interact with HP1  $\alpha$  and  $\beta$  *in vitro* (Figure 5.1(B)). This pentapeptide sequence (highlighted in green) is located adjacent to the first zinc coordinating cysteines of PHD1 within the DPF domain (highlighted in yellow). This region has rather poor electron density in the DPF/H3 crystal, although the sequence immediately adjacent to the PXVXL motif is an  $\alpha$ -helix (highlighted in blue).

189-PPVSLLPHEKDKPVAEPIPICSFC-212

Using site-directed mutagenesis, we have generated FLAG-MOZ constructs that contain a double substitution, P189A and V-191A, to test whether when mutated the MOZ-HP1 interaction is compromised. However, there was not sufficient time to complete this experiment before thesis submission.

Preliminary results indicate that MOZ is able to interact with all three HP1 isoforms. Further studies are required to reinforce the *in vivo* evidence and to map the region of MOZ responsible for the HP1 interaction. However, this data collectively for the first time shows that MOZ recruits proteins that sense H3K9 trimethylation.

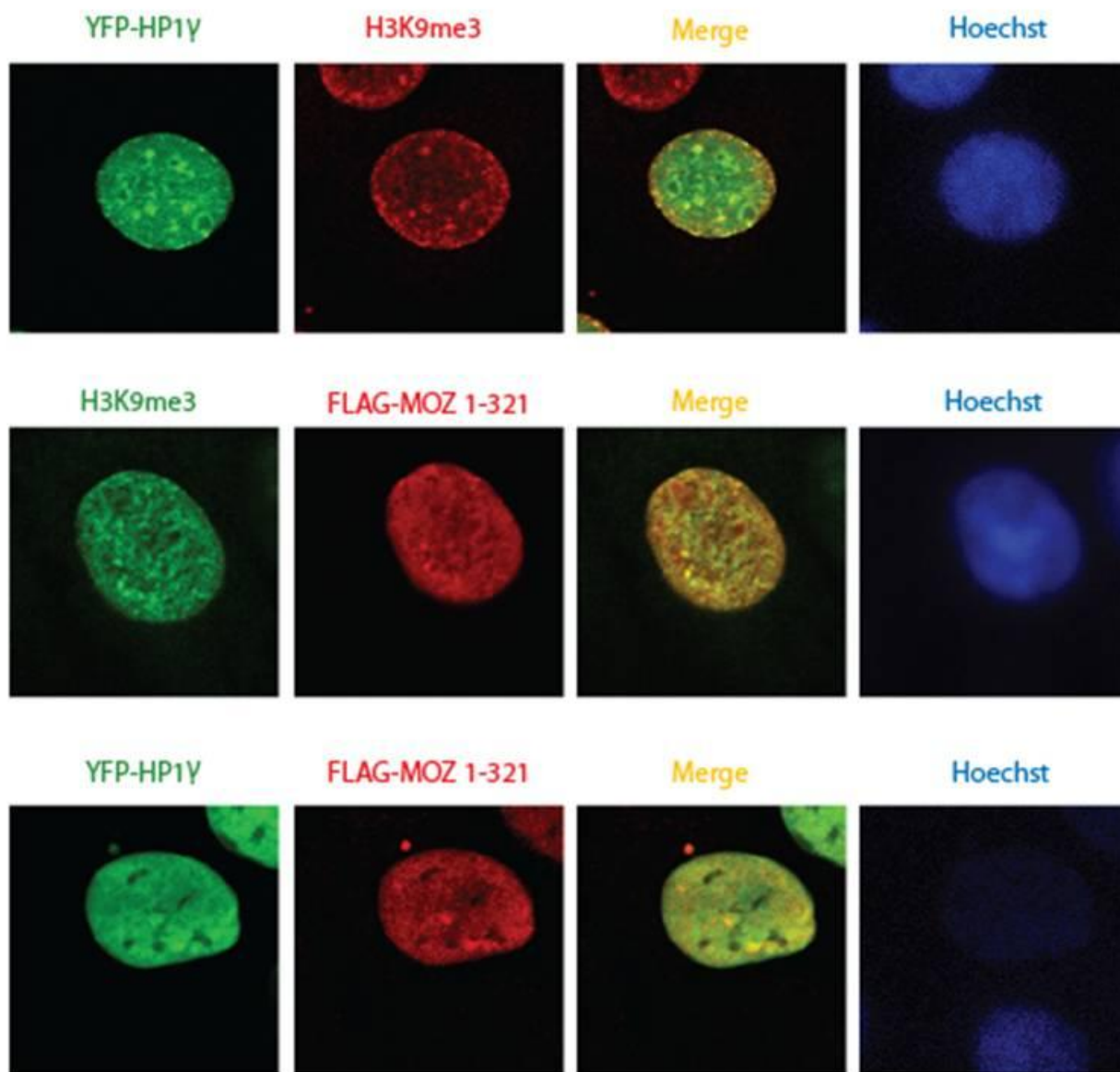
### 5.3. MOZ colocalises with HP1 $\gamma$

The previous data indicates that MOZ and HP1 $\gamma$  are associated in complexes *in vivo* and that this may be due to a PxV $\psi\psi$  motif adjacent to the MOZ DPF domain. To investigate this further, we next examined the subcellular localisation of exogenous FLAG-MOZ 1-321 and HP1 $\gamma$  proteins using immunofluorescence microscopy. As previously explained, HP1 $\alpha$  or HP1 $\beta$  could not be tested using this protocol due to the unavailability of suitable antibodies or epitope-tagged constructs.

To examine the subcellular localisation of MOZ 1-321 and HP1 $\gamma$ , U2OS cells were seeded onto coverslips and transfected with the FLAG-MOZ 1-321 and YFP-HP1 $\gamma$  expression plasmids. After 48 hours post-transfection cells were fixed, permeabilised and incubated with anti-FLAG antibody. YFP-HP1 $\gamma$  fluoresces directly so primary and secondary antibodies were not required to visualise the localisation of HP1 $\gamma$ . Both proteins were also tested for their association with the endogenous histone H3K9me3 PTM using the anti-H3K9me3 specific antibody. Incubation with either a mouse Alexa594-conjugated or a mixture of mouse Alexa594- and rabbit488-conjugated secondary antibodies followed. FLAG (red), HP1 $\gamma$  (green) and H3K9me3 (red/green - secondary used depended on the protein the PTM was being colocalised with) images were merged (yellow) to check for the nuclear localisation of the three

markers. The nucleus was visualised with Hoechst 33258 stain which intercalates with DNA (blue).

As shown in Figure 5.3, YFP-HP1 $\gamma$  was present in the nucleus of transfected U2OS cells. The distribution pattern showed that it is non-uniformly distributed in the nucleus, being excluded from nucleolar centres and concentrated in nucleolar ‘caps’ and other foci, as well as foci close to the nuclear periphery. These are presumably regions of heterochromatin. Staining for endogenous H3K9me3 revealed similar non-uniform, non-nucleolar distribution into nuclear foci. Merging of the images revealed at least partial colocalisation of HP1  $\gamma$  and H3K9me3, suggesting that some of these foci are heterochromatic regions (Figure 5.3- top panel). FLAG-MOZ 1-321 protein was also entirely nuclear suggesting that this region of MOZ contains a nuclear localisation sequence, or that association with chromatin or nuclear proteins keeps it localised to the nucleus. The distribution of FLAG-MOZ 1-321 was different from that of full-length MOZ (Collins et al., 2006; Kindle et al., 2010; Kindle et al., 2005) in that it was quite uniformly distributed throughout the nucleus. Full-length MOZ shows a non-nucleolar, speckled pattern much more similar to that of HP1 $\gamma$  (Kindle et al., 2005). Although partial localisation may be occurring between the N-terminus of MOZ and HP1 and H3K9me3 (Figure 5.3- middle and bottom panels), these data are insufficient to conclude whether MOZ and HP1 $\gamma$  are colocalised *in vivo*. However, it can be concluded here that other sequences in full length MOZ are likely to be required for its proper targeting to chromatin, especially heterochromatic regions.



**Figure 5.3. Colocalisation of MOZ with H3K9me3 and HP1 $\gamma$ .** U2OS cells were seeded onto coverslips and transfected after 24 hours with expression plasmids encoding FLAG-MOZ 1-321 and YFP-HP1 $\gamma$ . After 48 hours post-transfection, cells were fixed in 4 % paraformaldehyde, permeabilised and stained with antibodies raised against mouse anti-FLAG (red) and the rabbit anti-H3K9me3 (green). Staining was followed by incubation in anti-mouse Alexa594 and anti-rabbit Alexa488/594 secondary antibodies. Merged scans (yellow) highlight colocalisation and Hoechst 33258 (blue) was used as a stain of DNA.



#### 5.4. Differential effects of MOZ and HP1 $\gamma$ on AML1-mediated transcription

The *in vitro* evidence presented here suggests a potential interaction between MOZ and the three isoforms of HP1. Thus, to explore the functional consequences of the MOZ-HP1 interaction on the transcriptional activation properties of MOZ, luciferase reporter assays were conducted using the AML1-responsive promoter reporter system. It has been well-established that MOZ binds AML1 and strongly stimulates the transcription of AML1-dependent genes (Kitabayashi et al., 2001a), however despite a role for SUV39H1 in AML1 reporter repression (Chakraborty et al., 2003), a function for HP1 has yet to be investigated.

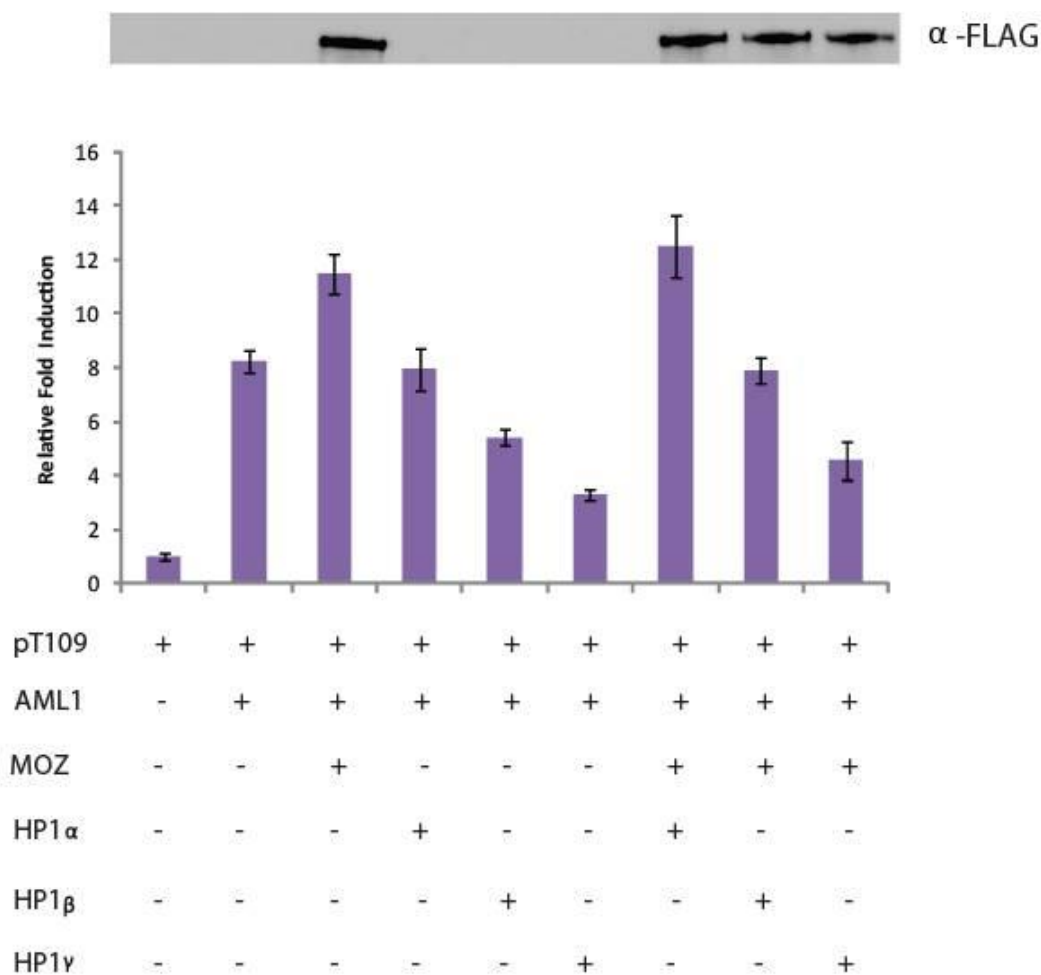
U2OS cells were transiently co-transfected with the pT109 3X-AML1 responsive luciferase reporter plasmid, a CMV- $\beta$ -galactosidase internal control reporter in addition to AML1, FLAG-MOZ and/or HP1 $\alpha$ , $\beta$ , $\gamma$  expression vectors as indicated. Luciferase activity was assayed 48 h post transfection using the Dual-Light Luciferase and  $\beta$ -Galactosidase reporter gene assay system (Applied Biosystems), measuring these activities using a Luminometer Orion (Berthold) as described in the Materials and Methods. The efficiency of FLAG-MOZ expression was checked by immunodetection using antibodies raised against the FLAG-tag (Figure 5.4). The expression of the HP1 proteins was not assessed due to previously described antibody issues, however HP1 $\gamma$  has already been shown to express efficiently in U2OS cells via other protocols, such as immunofluorescence.

Firstly, in agreement with previous reports from our group (Collins et al., 2006), AML1 activated the reporter approximately 8-fold over background and this

was further stimulated (to 12-fold) by co-expression with MOZ (Figure 5.4- compare bars 1 with 2 & 3). In contrast, both HP1 $\beta$  and HP1 $\gamma$  were found to repress the AML1-mediated reporter activity, resulting in 75% and 40% residual activity, respectively (Figure 5.4- compare bars 2 with 5 & 6). HP1 $\alpha$  did not appear to be able to act as a corepressor of AML1-driven transcription (Figure 5.4- compare bars 2 & 4). However, we were unable to confirm the expression of the HP1 proteins by western blots. Nonetheless, this result provides preliminary evidence that the HP1 $\gamma$  isoform acts as a repressor of AML1-dependent gene activity, although whether this affect is a consequence of an AML1-HP1 $\gamma$  interaction, the formation of an AML1-HP1 $\gamma$ -SUV39H1 co-repressor complex or via an alternative mechanism requires further investigation.

Having established that HP1 $\beta$  and  $\gamma$  repress AML1-mediated transcription, we next assessed what effect adding the coactivator MOZ and repressor HP1 simultaneously would have on reporter activation. As before, AML1-MOZ activated the reporter 12-fold above background and this was unaffected by coexpression of HP1 $\alpha$  (with the caveat that expression of HP1 $\alpha$  needs to be confirmed) (Figure 5.4- compare bars 3 & 7). The co-expression of MOZ and HP1 $\beta$  repressed the AML1 reporter activity to that observed in the absence of MOZ (Figure 5.4- compare bars 5 & 8), whereas reporter activation in the presence of both MOZ and HP1 $\gamma$  was only 4-fold above background (Figure 5.4- compare bars 6 & 9). These results suggest that MOZ and HP1 proteins, and in particular HP1 $\gamma$ , have opposing effects on the ability of AML1 to drive reporter gene transcription, acting as coactivators and corepressors respectively.

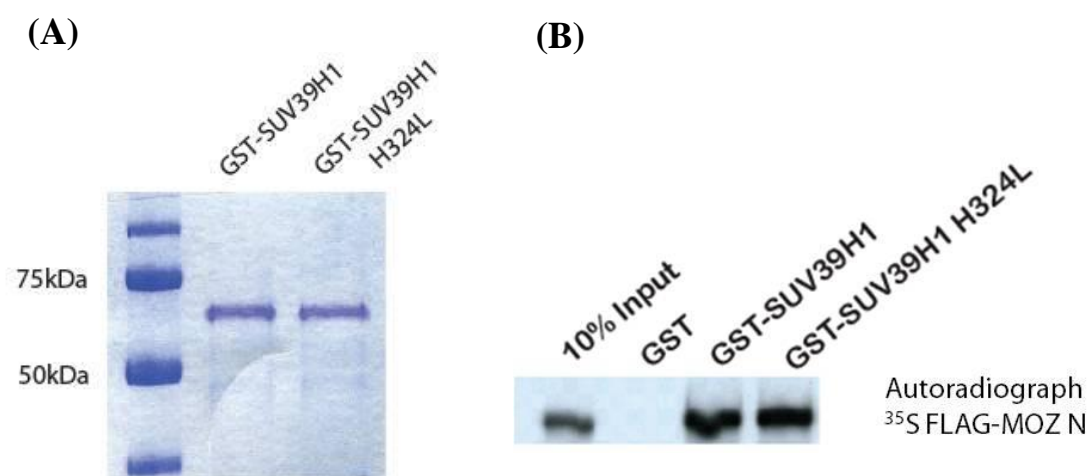
Interestingly, whereas HP1 $\alpha/\beta$  localise to pericentromeric heterochromatin, HP1 $\gamma$  is dispersed throughout the nucleus and associates with the coding regions of active genes (Maison and Almouzni, 2004). These results suggest two separate activities for MOZ at an active AML1-responsive promoter, possibly dependent on the associated proteins. Other groups have shown that MOZ is part of a possible coactivator complex, containing other HATs such as p300 and CBP, that was isolated by the affinity purification of AML1 (Yoshida and Kitabayashi, 2008). On the other hand, our data show that MOZ may also have a role in recruiting corepressors such as SUV39H1 (see later sections) and HP1 $\gamma$ , to repress AML1 target genes. This may be particularly important for understanding the action of oncogenic fusion proteins such as AML1-ETO, which appear to function principally as gene repressors. Further studies will be required to understand the possible role of MOZ in functional switches regulating gene expression.



**Figure 5.4. Over-expression of MOZ does not rescue HP1 $\gamma$ -mediated repression at the AML1 promoter.** U2OS cells were cotransfected as follows: with reporter plasmid pT109-3xAML1-luciferase (bar 1) and AML1 in the absence (bar 2) or presence of MOZ (bar 3), HP1 $\alpha$  (bar 4), HP1 $\beta$  (bar 5) or HP1 $\gamma$  (bar 6), MOZ & HP1 $\alpha$  (bar 7), MOZ & HP1 $\beta$  (bar 8) or MOZ & HP1 $\gamma$  (bar 9). A plasmid encoding  $\beta$ -galactosidase was used to normalise the transfection efficiency. Reporter activation is represented as a fold induction over the control value (reporters in the absence of coactivators). The data are presented as averages and standard errors of the means from triplicate samples.

## 5.5. MOZ and SUV39H1 interact *in vitro*

Given that MOZ can associate with histone H3 tri-methylated K9 and HP1 proteins, we asked whether it can also form a functional complex with SUV39H1. GST-SUV39H1 WT and mutant fusion proteins were expressed and purified as described in previous sections (sections 3.4 & 4.2). Protein concentrations were normalised by SDS-PAGE and Coomassie staining to ensure equal amounts of protein were used in subsequent *in vitro* assays (Figure 5.5(A)). To assess whether MOZ and SUV39H1 interact with one another *in vitro*, GST pulldown assays were conducted using WT GST-SUV39H1 or GST-SUV39H1 containing the mutation H324L (a point mutant that renders the SET methyltransferase domain inactive) and FLAG-MOZ N protein that had been *in vitro* translated in the presence of <sup>35</sup>S-methionine. The GST-SUV39H1 constructs were provided as gifts by Dr. Thomas Jenuwein, whereas FLAG-MOZ N was described previously (Kindle et al., 2005). The input lane shows efficient IVT of the MOZ protein (Figure 5.5(B)-lane 1). As a control reaction, MOZ was incubated with the GST protein alone. No interaction was observed between the GST-conjugated beads and MOZ, proving an interaction does not occur through the GST tag (Figure 5.5(B)-lane 2). MOZ was however pulled down with both GST-SUV39H1 WT and H324L-conjugated beads (Figure 5.5(B)-lanes 3&4). This result shows that MOZ can interact with SUV39H1 *in vitro*, irrespective of its methyltransferase activity.



**Figure 5.5. MOZ interacts with the lysine methyltransferase SUV39H1 *in vitro*.** (A) Normalisation of GST-SUV39H1 constructs concentration by SDS-PAGE and Coomassie staining. (B) Pulldown assays with 1.5 µg/32 nM GST-SUV39H1 and *in vitro*-translated <sup>35</sup>S-methionine labelled FLAG-MOZ N followed by autoradiograph analysis.

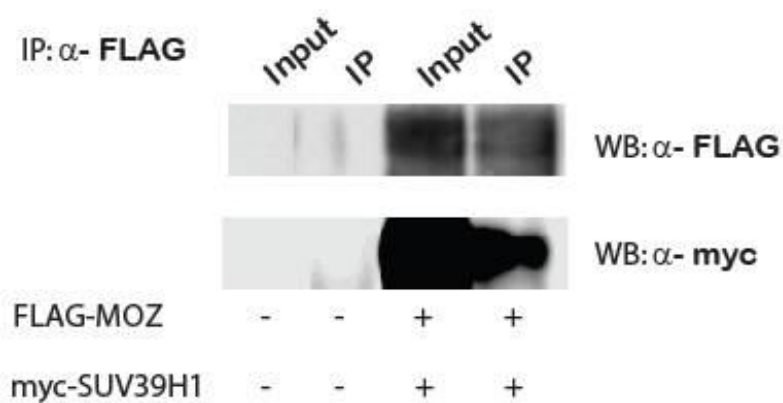
## 5.6. MOZ interacts with SUV39H1 *in vivo*

To validate that MOZ and SUV39H1 can interact with one another in mammalian cells, co-immunoprecipitations were performed with FLAG-MOZ and myc-SUV39H1 proteins expressed in HEK293 cells. The myc-SUV39H1 constructs were provided as gifts from Dr. Thomas Jenuwein. FLAG-MOZ and myc-SUV39H1 exogenous proteins were transiently transfected into HEK293 cells and 48 hours post-transfection, cells were harvested and lysed to release the soluble proteins. Input lanes verify the expression of each transiently transfected protein in the cell lysates (Figure 5.6(A)- lanes 1 & 3 and (B)- lanes 1 & 2). FLAG-MOZ proteins were then immunoprecipitated from the cell lysates with a FLAG antibody and tested for the SUV39H1 interacting partner by immunodetection with the anti-Myc antibody (Figure 5.6). To determine the efficiency of the immunoprecipitation, blots were also probed with the FLAG antibody. As expected, the FLAG antibody immunoprecipitated a band at ~220kDa (Figure 5.6(A)-top panel), only in lanes that expressed FLAG-MOZ. A protein at ~50kDa was detected when both FLAG-MOZ (full-length) and myc-SUV39H1 were co-expressed, but not in control mock transfected lysates (Figure 5.6(A)- bottom panel, compare lanes 2&4). The identity of this protein was confirmed using the antibody recognising the myc tag. This result confirms that MOZ and the SUV39H1 methyltransferase interact *in vivo* as well as *in vitro*. Co-immunoprecipitation was also carried out with exogenously expressed FLAG-MOZ 1-321. Western blot analysis using FLAG antibody showed immunoprecipitation of a ~37kDa (Figure 5.6(B)-bottom panel) band only in lanes that expressed FLAG-MOZ 1-321. Subsequent blotting using the myc antibody

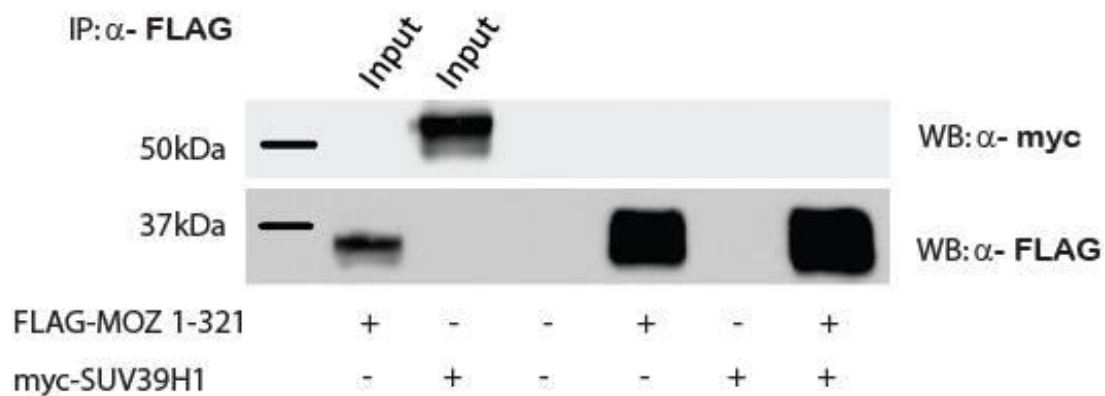
revealed the detection of no protein around the 50kDa molecular weight when both proteins were co-expressed (Figure 5.6(B)-top panel, lane 6), despite the availability of SUV39H1 in the input lysate (Figure 5.6(B)-top panel, lane 2). This indicates that myc-SUV39H1 was not co-immunoprecipitated with FLAG-MOZ 1-321, therefore the interaction with SUV39H1 does not appear to be mediated via the DPF or other sequences in the N-terminus of MOZ.



(A)



(B)

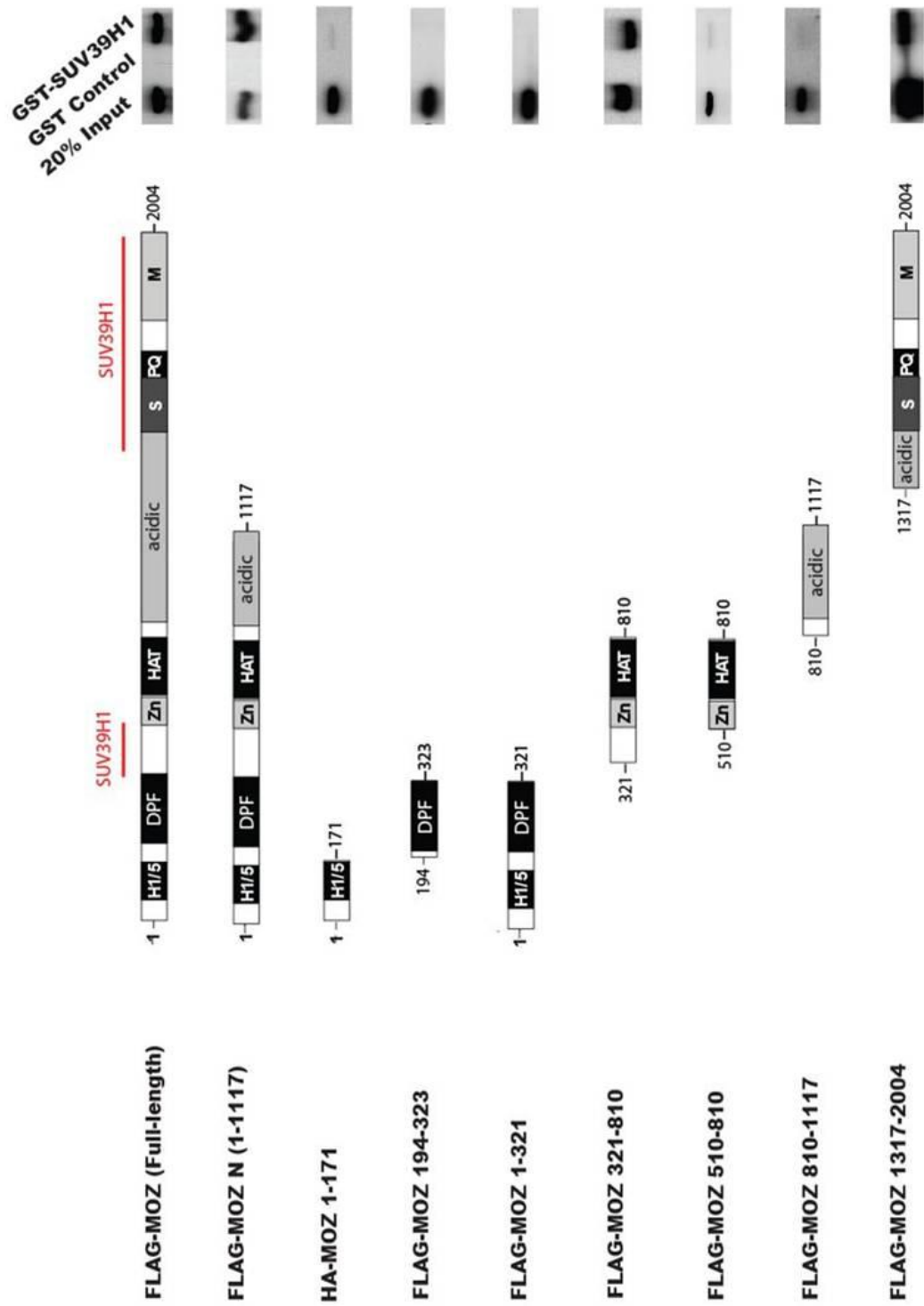


**Figure 5.6. MOZ interacts with the lysine methyltransferase SUV39H1 *in vivo*.** HEK293 cells were mock transfected (lanes (A) 1 & 2 and (B) 3) or transiently cotransfected with Myc-SUV39H1 and (A) FLAG-MOZ (lanes 3 & 4) or (B) FLAG-MOZ 1-321 (lanes 4-6). 48 h post-transfection cells were lysed and the proteins were coimmunoprecipitated with (A)  $\alpha$ -MOZ or (B)  $\alpha$ -FLAG. Proteins were separated by electrophoresis and subjected to immunoblotting using (A)  $\alpha$ -MOZ or (B)  $\alpha$ -FLAG and  $\alpha$ -myc antibodies.

## 5.7. SUV39H1 interacts with two regions of MOZ

Having observed an interaction between MOZ and SUV39H1 both *in vitro* and *in vivo*, we subsequently set out to identify the region of MOZ responsible for the interaction. To map the MOZ-SUV39H1 interaction, a panel of FLAG-MOZ constructs covering all regions of MOZ was generated. The constructs FLAG-MOZ 1-171, 1-321, MOZ N, MOZ (full-length) were described previously (Collins et al., 2006; Kindle et al., 2005). In addition, FLAG-MOZ 194-323, 321-810, 510-810, 810-1117 and 1317-2004 were generated by PCR subcloning. The primers are listed in appendix table A.1. All constructs were verified by PCR and sequence analysis. The FLAG-MOZ constructs were *in vitro* translated and incubated in a GST pulldown assay with either the GST control protein or GST-SUV39H1. Input lanes show the efficient *in vitro* translation of each of the MOZ proteins (Figure 5.7-lane 1). No interaction was observed between the GST control protein and MOZ (Figure 5.7-lane 2). Confirming previous results, both the full-length and MOZ N proteins were able to interact with the methyltransferase (Figure 5.7- panels 1 & 2), suggesting binding of SUV39H1 does not require the C-terminal half of the MOZ protein. Consequently, a variety of N-terminal constructs were tested for interaction with SUV39H1 (Figure 5.7- panels 3-5). Results indicate that both the H1/5 and DPF modules individually (panel 3-4 respectively) and together (panel 5) are unable to associate with the methyltransferase. A faint band can be seen for MOZ 1-171 (panel 3), which calculated by densitometry, represents 0.8 % total binding relative to control. However, due to a very weak signal and no interaction established for MOZ 1-321 (panel 5) this was discarded as background binding. This data also corroborates

previous *in vivo* results, whereby SUV39H1 was unable to coimmunoprecipitate with the MOZ N-terminus (amino acids 1-321) (Figure 5.6(B)). Next, constructs representing internal MOZ protein sequences were tested for interaction (Figure 5.7-panels 6-8). The MOZ 321-810 construct, which contains sequence of unknown function and the MYST domain, showed a robust association with SUV39H1 (panel 6). However, the MYST domain alone (510-810) did not interact with SUV39H1 (panel 7). This suggests that the linker region between the DPF and MYST domains, corresponding to amino acids 321-510, constitutes a SUV39H1 binding site. The N-terminal region of the acidic domain, amino acids 810-1117, also failed to bind to SUV39H1 (panel 8). To explore whether the C-terminus of MOZ has any role in SUV39H1 interaction we finally tested FLAG-MOZ 1317-2004, containing the acidic and S/M-rich region, in the GST pulldown assays. This identified a second strong SUV39H1 binding site in the MOZ C-terminus (panel 9) and indicates that the methyltransferase interacts with two regions of the MOZ protein (summarised in red above the full-length schematic). Intriguingly, the SUV39H1 binding sites identified here overlap with those identified previously for the MOZ/AML1 interaction, whereby MOZ recruits AML1 via two sites mapped to 312-664 and 1517-2004 (Kitabayashi et al., 2001a). This raises the possibility that SUV39H1 and AML1 may be recruited to similar regions of the MOZ protein, or alternatively might indicate competitive binding to MOZ. There was not sufficient time to pursue this hypothesis.



**Figure 5.7. SUV39H1 interacts with two regions of MOZ.** Schematic representation of the structure of MOZ deletion constructs used for determining the SUV39H1-interacting domains (left panel). Pull-down assays were carried out with *in vitro*-translated <sup>35</sup>S-methionine labelled FLAG-MOZ deletion constructs and 1.5 µg/ 32 nM GST-SUV39H1 and analysed by autoradiography (right panel). Regions of MOZ required for interaction with SUV39H1 are highlighted in red on the full-length schematic.

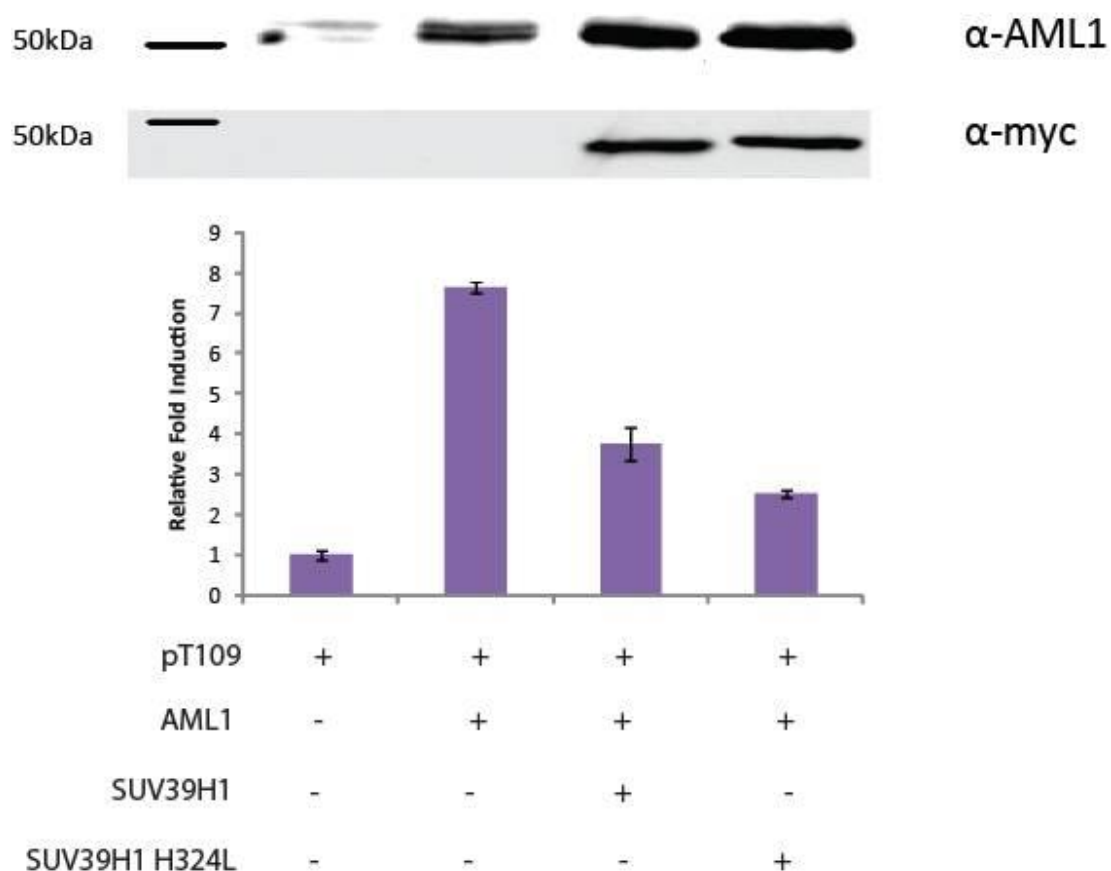
## 5.8. MOZ does not rescue SUV39H1-mediated repression of AML1 transactivation

To investigate the functional consequences of the MOZ-SUV39H1 interaction on the transcriptional activation properties of MOZ, luciferase reporter assays were conducted utilising an AML1 responsive promoter/reporter system. Previous literature shows that MOZ interacts with AML1 and strongly stimulates transcription of AML1-dependent genes (Kitabayashi et al., 2001a) yet, conversely, SUV39H1 also forms a stable complex with AML1 resulting in the abrogation of DNA-binding and transactivation properties of AML1 (Chakraborty et al., 2003). Thus, to explore the opposing effects of these cofactors on AML1 function, we performed reporter assays.

U2OS cells were transiently co-transfected with the pT109 3X-AML1 responsive luciferase reporter plasmid, a CMV- $\beta$ -galactosidase internal control reporter, in addition to AML1, FLAG-MOZ and/or myc-SUV39H1 expression vectors as indicated. Luciferase activity was assayed 48 h post transfection using the Dual-light Luciferase and  $\beta$ -Galactosidase reporter gene assay system (Applied Biosystems), measuring these activities using a Luminometer Orion (Berthold) as described in the Materials and Methods. Firstly, we assessed the effect of SUV39H1 WT or the SET domain mutant SUV39H1 H324L on AML1 driven reporter activity. Expression levels of exogenous proteins were assessed by immunodetection using antibodies raised against the myc tag and the AML1 protein itself (Figure 5.8). The western blots showed that both proteins were expressed at similar levels only in cell lysates that were transfected with that given protein. As shown in Figure 5.8, AML1

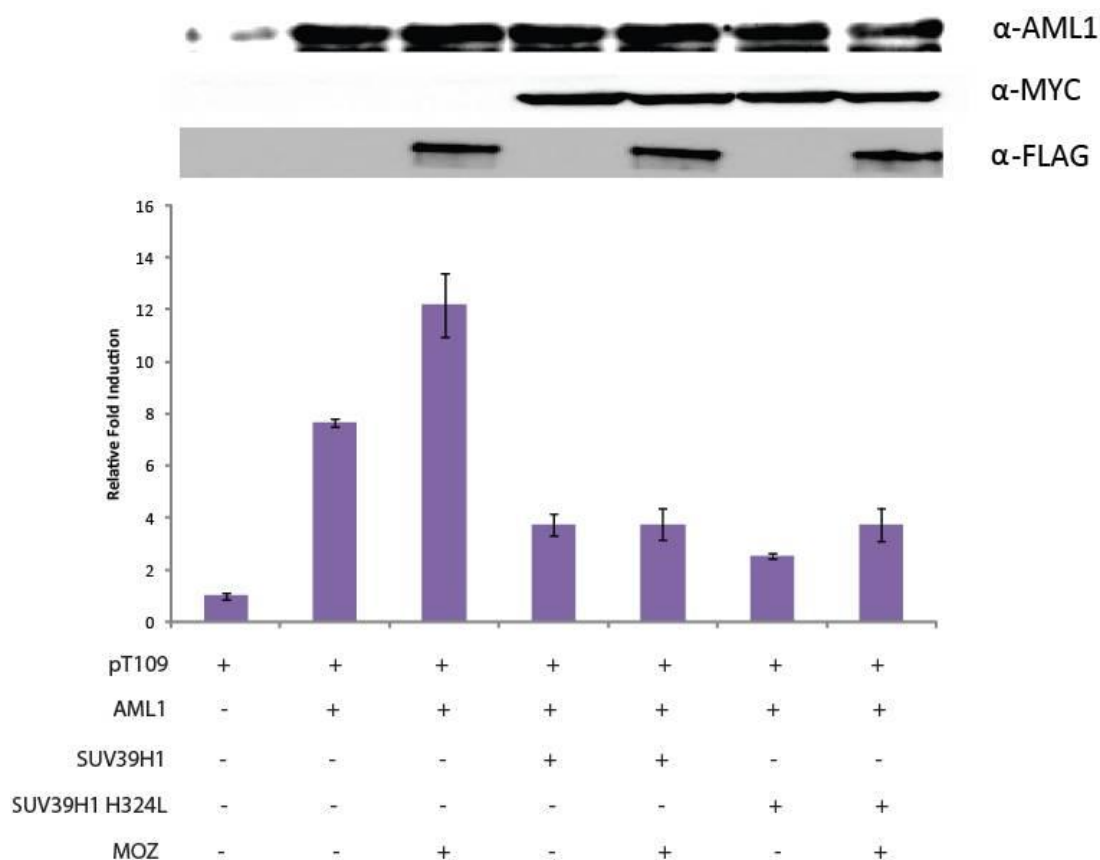
activated the reporter approximately 8-fold over the basal activity, consistent with previous reports. Co-expression of SUV39H1 reduced AML1-mediated reporter activation in a dose-dependent manner (Figure 5.8- compare bars 2 & 3). A similar repressive activity was observed using the H324L mutant, indicating that the methyltransferase activity of SUV39H1 is not essential for repression of the AML1 reporter (Figure 5.8- compare bars 2 & 4). In agreement with our previous data (Collins et al., 2006), co-expression of AML1 with MOZ further enhanced the reporter activation to 12-fold above basal level (Figure 5.9- compare bars 2 & 3). Expression of FLAG-MOZ was also verified by western blots (Figure 5.9). Interestingly, when MOZ and SUV39H1 were expressed together with AML1, the transactivation property of AML1 was still impaired (Figure 5.9-compare bars 3 & 4 with 5), both by the WT and mutant SUV39H1 proteins (Figure 5.9- compare bars 3 & 6 with 7). Thus, MOZ cannot rescue SUV39H1-mediated repression at an AML1 targeted promoter. This suggests that the repressive activities of SUV39H1 dominate over the stimulating transactivation properties of MOZ at the AML1-responsive promoter. These results combined with *in vitro* mapping data suggest a possible model where SUV39H1 sequesters the MOZ protein by blocking the MOZ-AML1 interaction regions. *In vitro* data show a clear overlapping of the two regions required for AML1 (312-664 & 1517-2004) and SUV39H1 (321-510 & 1317-2004) binding, implying that only one of the proteins can bind at any given time. Interaction with SUV39H1 might therefore prevent MOZ from binding the AML1 protein, allowing free access for the methyltransferase to act as a corepressor at the AML1-responsive reporter. Of course this mechanism of SUV39H1 dominance over MOZ may only occur at the AML1-responsive promoter and in order to test this theory the reporter

activity should be monitored at other MOZ driven target genes, such as the *Hox* genes.



**Figure 5.8. SUV39H1 represses the transcriptional activity of AML1.** U2OS cells were cotransfected as follows: with reporter plasmid pT109-3x AML1-luciferase (bar 1) and AML1 in the absence (bar 2) or presence of SUV39H1 (bar 3) or an inactive point mutant SUV39H1 (bar 4). A plasmid encoding  $\beta$ -galactosidase was used to normalise the transfection efficiency. Reporter activation is represented as a fold induction over the control value (reporters in the absence of coactivators). The data are presented as averages and standard errors of the means from triplicate samples.





**Figure 5.9. Over-expression of MOZ does not rescue SUV39H1-mediated repression at the AML1 promoter.** U2OS cells were cotransfected as follows: with reporter plasmid pT109-3x AML1-luciferase (bar 1) and AML1 in the absence (bar 2) or presence of MOZ (bar 3), SUV39H1 (bar 4) an inactive point mutant SUV39H1 (H324L) (bar 6), MOZ & SUV39H1 (bar 5) or MOZ & SUV39H1 H324L (bar 7). A plasmid encoding  $\beta$ -galactosidase was used to normalise the transfection efficiency. Reporter activation is represented as a fold induction over the control value (reporters in the absence of coactivators). The data are presented as averages and standard errors of the means from triplicate samples.

**5.9. Summary**

Studies in the literature, including genome wide ChIP mapping, have revealed that H3K4me3 generally correlates with transcriptionally active genes, whereas, H3K9me3 is associated with the formation of heterochromatic, transcriptionally inert regions (Kouzarides, 2007). In chapter three, MOZ was shown to readily associate with chromatin enriched in trimethylated H3K9, but to be excluded from chromatin containing H3K4me3. These binding properties were shown to be determined by the DPF. These findings can be interpreted in a number of ways with respect to the context of understanding the role of MOZ as a co-regulator of gene transcription. MOZ has been suggested to function as a coactivator, such as for the Runx family of transcription factors including AML1. Recruitment of MOZ to chromatin containing repressive marks such as H3K9me3 may indicate that MOZ functions at an early stage of the gene activation process. The ability of MOZ to associate with heterochromatic PTMs, may allow it to induce the removal of repressive PTMs and catalyse the deposition of activating PTMs such as H3K9 and H3K14 acetylation. This in turn may lead to the recruitment of other factors such as the MLL/WDR5 complex which signal the exit of MOZ from chromatin, e.g. as directed by H3K4me3 allowing it to process neighbouring heterochromatic regions, or be recycled to other gene promoters. MOZ may stimulate this itself through the acetylation of H4, a second exit signal recognised by the DPF.

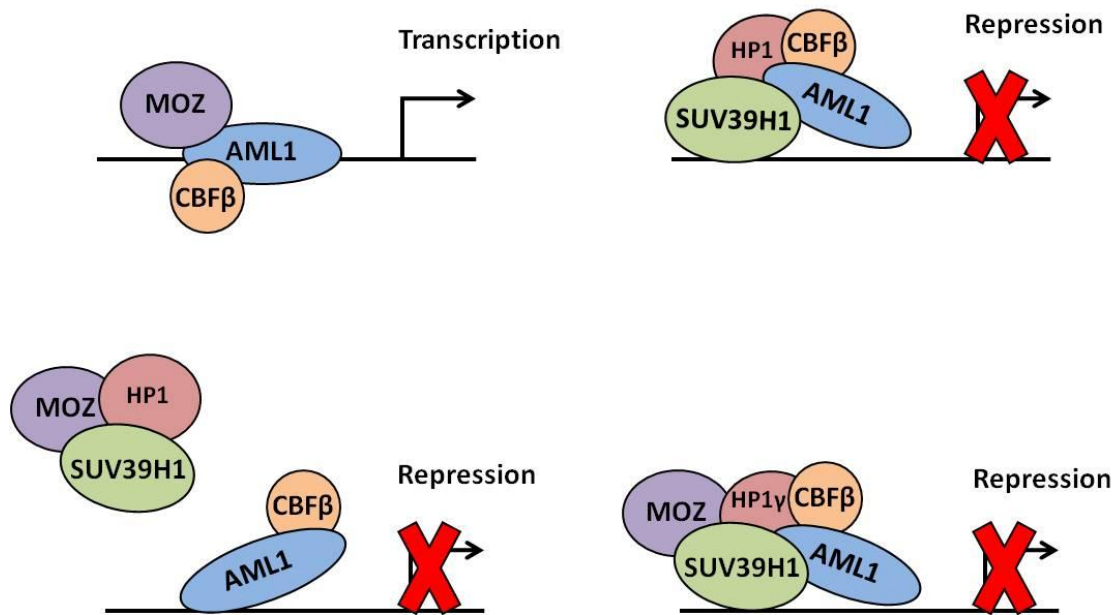
Although protein interaction studies will not address the sequence of events that occur in realtime during the activation process, they will be useful in shedding light on the nature of molecular interactions that can occur and how these might be

regulated by posttranslational modifications. We therefore undertook binding studies to try to better understand the relationship between MOZ and H3K9me3. SUV39H1 is the methyltransferase which catalyses the trimethylation of H3K9 that consequently serves as a binding site for HP1. Thus, we began by asking whether MOZ could associate with such proteins that impart and sense the trimethylation of H3K9.

Using GST pulldown assays and coimmunoprecipitation experiments, it was demonstrated that MOZ and SUV39H1 could associate directly with each other, and associate in complexes both *in vitro* and *in vivo*. This ability to interact was independent of the methyltransferase activity of SUV39H1, as both WT and a SET domain point mutant were competent to bind MOZ. Mapping studies excluded a role for the H1/5 and DPF domains in SUV39H1 binding and identified two regions that appear to be independently capable of SUV39H1 binding. Immunofluorescence studies were hampered by poor quality of the commercial myc-antibody. To progress with this study, an epitope-tagged version of SUV39H1 will need to be generated.

The two distinct regions of MOZ that bind SUV39H1 were mapped to the sequence between the DPF and MYST domains, (amino acids 321-510) and a C-terminal fragment of MOZ, containing the acidic and S/M-rich region (1317-2004). Interestingly, these sites overlap with binding sites identified for AML1 and p53 protein interactions, mapped to the linker region between DPF and MYST (312-664) and the C-terminus (1517-2004) (Kitabayashi et al., 2001a; Rokudai et al., 2009). MOZ functions as a transcriptional coactivator for AML1-mediated transcription (Kitabayashi et al., 2001a), whereas AML1 interaction with SUV39H1 leads to the repression of AML1-dependent target genes (Chakraborty et al., 2003). Thus, both

MOZ and SUV39H1 function as coregulators of AML1-dependent gene expression. Therefore, we next asked how the MOZ-SUV39H1 interaction affects the transactivation properties at an AML1-dependent promoter. Reporter assays utilising an AML1-responsive reporter system in U2OS mammalian cells revealed that MOZ activates AML1-mediated expression of the reporter, whereas SUV39H1 repressed the reporter activity. Co-expression of MOZ and SUV39H1 also resulted in the repression of the reporter. Although the result may be influenced by the relative expression levels of both proteins, it also indicates that MOZ is unable to prevent the repression by SUV39H1 under these conditions. The methyltransferase activity of SUV39H1 was not required for the repression in these assays. This indicated that repression was not enforced by the post-translational methylation of either AML1 or MOZ by SUV39H1 itself. The reporter assay results, in association with *in vitro* mapping data, suggest a model where SUV39H1 competitively blocks the interaction between MOZ and AML1, preventing the transcriptional upregulation of AML1-dependent genes (Figure 5.10).



**Figure 5.10. Schematic model illustrating possible events at the AML1 reporter.** SUV39H1 competitively blocks the MOZ-AML1 binding site, displacing or sequestering MOZ from the AML1 complex and preventing the activation of the AML1-dependent reporter.

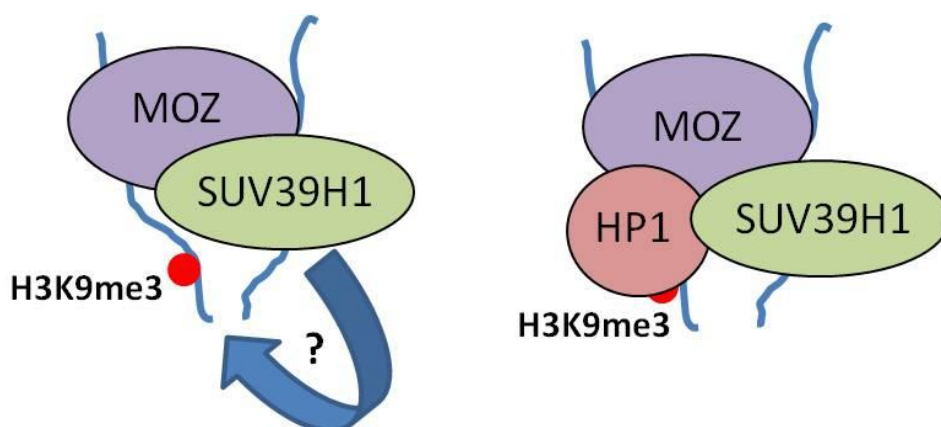
SUV39H1 interacts and functions in complex with HP1, thus we next investigated whether MOZ could also interact with the HP1 protein. GST-pulldown assays revealed an interaction between the N-terminus of MOZ (1-321) and HP1 $\alpha$  and  $\beta$  and coimmunoprecipitation experiments showed that MOZ 1-321 could directly associate with HP1 $\gamma$ . *In vivo* analysis of the HP1 $\alpha$  and  $\beta$  isoforms were hindered by the poor quality of the commercial HP1-antibody, thus to progress with this study, an epitope-tagged version of HP1 $\alpha$  and  $\beta$  will need to be generated. However, preliminary results indicated that the N-terminus of MOZ is able to interact with all three isoforms of HP1.

Analysis of the MOZ N-terminal sequence highlighted the presence of a conserved pentameric motif (PxV $\psi\psi$  - where  $\psi$  represents a hydrophobic residue), which may be sufficient for specific interaction with the HP1 chromoshadow domain (Cowieson et al., 2000). The disruption of this motif in MOZ by site-directed mutagenesis will be required to test whether, when mutated, the MOZ-HP1 interaction is compromised. Of course, multiple binding sites for HP1 cannot be ruled out at this stage, as the full-length protein requires testing in all binding assays and the entire sequence needs to be analysed to predict other possible HP1-binding sites.

SUV39H1 has been shown to act as a corepressor at AML-responsive promoters, however whether HP1 also functions in the repression of AML1-dependent genes remains unclear. Therefore, we next asked whether the MOZ-HP1 interaction affects the transactivation properties at the AML1 promoter. Preliminary assays utilising an AML1-responsive reporter system in U2OS mammalian cells demonstrated that the HP1 $\beta$  and  $\gamma$  isoforms acted as repressors of AML1-mediated transcription, whereas HP1 $\alpha$  had no effect on AML1 transactivation. However, expression of the HP1 $\alpha,\beta,\gamma$  isoforms needs to be confirmed. Coexpression of MOZ and HP1 $\beta$  and, in particular, MOZ and HP1 $\gamma$  also resulted in repression of the reporter; results reminiscent of the affect observed for SUV39H1 at the AML1-responsive reporter. Although the result may be influenced by relative expression levels of each of the proteins, these results indicate that MOZ is unable to prevent the repression by HP1 under these conditions.

Interestingly, whereas HP1 $\alpha/\beta$  localise to pericentromeric heterochromatin, HP1 $\gamma$  is dispersed throughout the nucleus and associates with the coding regions of active genes (Maison and Almouzni, 2004). Little is known of the specific

biochemical mechanism by which HP1 inactivates a gene. However, it is understood that a promoter-proximal localisation of HP1 triggers silencing, concurrent with localised chromatin condensation and an increase in H3K9me3 (Maison and Almouzni, 2004). Moreover, it was discovered that HP1-mediated silencing targets RNAPII coactivator complexes, interacting with primary PIC components and inhibiting the assembly of the PIC at gene promoters (Smallwood et al., 2008). The above results suggested that as well as targeting RNAPII coactivator complexes, HP1 $\gamma$  (possibly in coordination with SUV39H1) may also interact with and inhibit the activities of regulatory complexes, such as the MOZ-AML1 complex, at active gene promoters. This would suggest a model whereby MOZ is able function as part of both a coactivator and corepressor complex. Firstly, MOZ may function as a coactivator during the early stages of the gene activation process, inducing the removal of repressive PTMs and catalysing the deposition of activating PTMs such as H3K9ac. Evidence now suggests that MOZ may also act to recruit other chromatin modulators to facilitate a repressive chromatin environment. Our data suggests this is through the recruitment of SUV39H1 by MOZ to the active coding region of genes, which consequently induces the trimethylation of H3K9, HP1 binding and the initiation of gene silencing (Figure 5.11).



**Figure 5.11. Model showing the hypothesised recruitment of transcriptional modulators by MOZ to chromatin.**

To investigate further into the existence of a MOZ/SUV39H1/HP1 corepressor complex, it would be important to utilise protocols that could validate these interactions in a promoter context. Thus, ChIP analysis could be used to check the occupancy of MOZ and SUV39H1 at AML1-dependent gene promoters, such as myeloperoxidase or at *Hox* gene loci. However, as the targets of MOZ and SUV39H1/HP1 are yet to be completely defined, ChIP Seq would be the most informative approach. This would allow the identification of the genome-wide binding sites associated with MOZ/SUV39H1/HP1 and enable mapping of their proximity to promoters and TSS regions. ChIP seq could also be used to correlate these binding sites to chromatin PTMs, such as H3K9me3. Interestingly, recent data from our group shows that MOZ expression is up-regulated during DNA damage, thus this novel corepressor complex may be induced in response to cellular genotoxic stress to halt transcription until DNA repair has prevailed. Indeed, the p53-MOZ



complex has been shown to function in response to genotoxic stress (Rokudai et al., 2009). Affinity purification of MOZ complexes from cells subjected to various DNA damage reagents could be carried out to identify if SUV39H1 and HP1 are present.

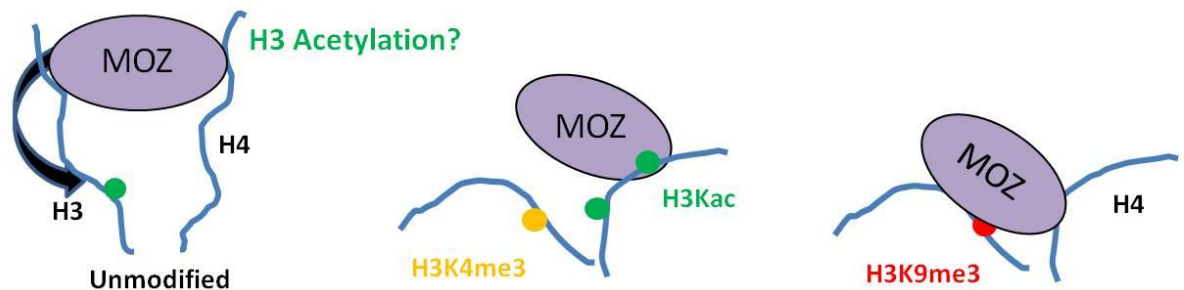
## **Chapter Six: Concluding Remarks**

MOZ is a histone acetyltransferase that functions as an essential coactivator for genes that are involved in haematopoietic and developmental processes. The MOZ MYST domain contains acetyltransferase activity towards histone H4 *in vitro* (Champagne et al., 2001). However, much remains to be discovered regarding how chromatin regulators, like MOZ, deposit PTMs onto chromatin and how these signals are ‘read’ combinatorially. Chromatin binding is a widespread property of PHD fingers, thus it was predicted that the N-terminal double PHD fingers of MOZ may function in targeting the enzyme to chromatin. In this study, we have shown that the MOZ DPF can sense selected acetyl and methyl PTMs in H3 and H4 and that the integrity of both PHD fingers was required for these interactions. Interactions with the DPF occurred via H3 (1-21) and H4 (2-24) N-terminal tails. Whereas dimethylation of H3K4 had no effect on binding, it was demonstrated that H3K4me3 strongly reduced interaction with MOZ. Similarly, H4 acetylation (K5,8,12,16) abrogated histone binding (Figure 6.1). In contrast, other PTMs such as H3R2me2a, H3K9me3 and H3K9/14ac did not adversely affect MOZ recruitment to histones or native chromatin. Furthermore, we showed that the separate PHD1 and PHD2 entities of the DPF displayed differential histone binding properties and functionally cooperate to engage chromatin PTMs bivalently; in particular to the discrimination of H3K4 trimethylation.

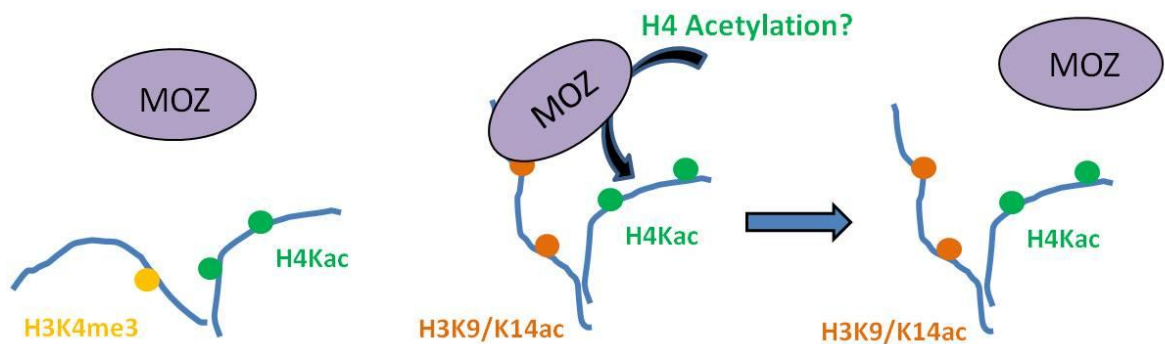
It is emerging that chromatin regulators can sense multiple PTMs on different histones simultaneously (Ruthenburg et al., 2007b). In our studies we showed that the availability of unmodified H3 or H4 histone tails rescued the lack of DPF association to H3K4me3, suggesting that H3K4me3 acts as a recessive exit signal that does not affect the simultaneous binding to PTM residues on neighbouring histone tails.

Conversely, H4Kac exerted a dominant inhibitory affect, preventing binding to unmodified H3, H3Kac and H3K9me3 histone tails (Figure 6.1). Quantitative experiments revealed a weak affinity of H4K16ac for MOZ DPF, indicating that H4Kac tail binding might induce a conformational change to displace H3 histone tails.

### RECRUITMENT SIGNALS



### RELEASE SIGNALS



**Figure 6.1. Schematic model depicting the recruitment and release signals for MOZ to and from chromatin.** MOZ recruitment is facilitated by binding to unmodified H3/H4 tails and is further stabilised by the acetylation of H3K9/14 by DPF+MYST. In addition, MOZ is targeted to chromatin enriched in H3Kac and H3K9me3 PTMs. Release signals include the presence of H3K4me3 and the induced or pre-written H4Kac histone PTMs.

## Chapter Six: Concluding Remarks

The MOZ DPF motif is shared by only a few other PHD-containing proteins, including the DPF protein family (Lange et al., 2008). A high conservation is shared between MOZ and DPF2/3 sequences and our studies revealed that the overall structural fold is conserved, but differences are apparent in the loop regions of the structure. However, structural analysis revealed a striking difference between the H3 peptide fold in the two complexes. The H3 peptide in the MOZ complex structure adopted an  $\alpha$ -helical conformation, compared to an unstructured conformation in the DPF3b complex; this to our knowledge has not been previously observed for the H3-terminus in interactions with PHD domains.

Recently, the interplay between histone ‘reader’ and ‘writer’ modules has been documented in the literature, including the synergy between the bromodomain and HAT domain of the acetyltransferase p300 (Chen et al., 2010). Remarkably, our preliminary data shows that the DPF domain can function as a mild H3-specific acetyltransferase. Neither PHD1 nor PHD2 alone could acetylate histones, thus indicating that the composite DPF is required for activity. These studies provide the first report of a possible enzymatic role in chromatin modification attributed to a PHD finger. The DPF was found to work in conjunction with the H4-specific MYST domain to direct histone acetylation and to extend the range of histone substrates targeted by MOZ. Moreover, the DPF domain was found to enhance the kinetics of MOZ-induced histone acetylation. We have also shown that trimethylation of H3K4 and monoacetylation of H4, prevented further acetylation of the H3 and H4 peptide respectively by MOZ. Conversely, the acetylation of H3 by MOZ DPF or DPF+MYST promotes binding to the H3 N-terminus, suggesting that MOZ may

promote its own recruitment to chromatin via acetylation of and binding to H3 (Figure 6.1).

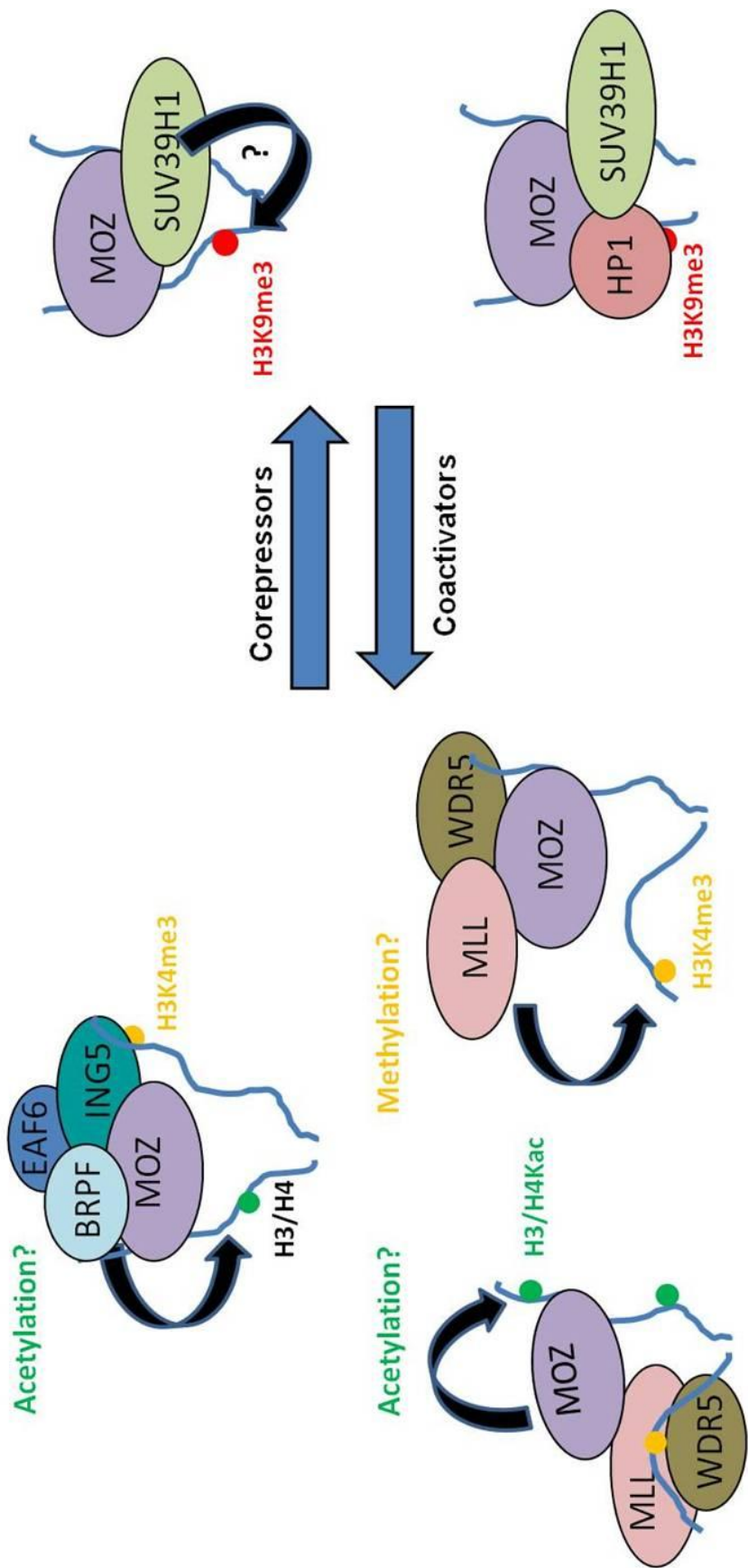
Our studies revealed that the MOZ DPF could associate with histones and native chromatin enriched in H3K9me3. H3K9me3 is associated with the formation of heterochromatin and transcriptionally inert regions (Kouzarides, 2007), thus we carried out studies to better understand the relationship between MOZ and H3K9me3. In this study we provide evidence demonstrating that MOZ could interact with the H3K9me3-specific methyltransferase, SUV39H1 and the H3K9me3-binding protein HP1. The two distinct regions of MOZ that bind SUV39H1 were similar to the mapped regions that have been established for AML1 (Kitabayashi et al., 2001a). Furthermore, it was shown that MOZ was unable to prevent SUV39H1- and HP1-mediated repression at the AML1-dependent reporter. This suggested a model whereby SUV39H1 competitively blocks the interaction between MOZ and AML1, preventing the upregulation of AML1-dependent genes.

Multicomponent nuclear machines are likely to use multivalent interactions with histone PTMs to engage chromatin and to install sequential or opposing signals leading to gene activation or repression. Thus, to provide insight into MOZ function and the sequence of events that takes place, further probing is required into the complexes in which MOZ resides. Based on our data and previous reports in the literature we propose the following model of events that influence the dynamic association of MOZ to chromatin (Figure 6.2). In this study we found that although MOZ DPF itself could not bind H3K4me3, this modification did not hinder its subsequent binding and acetylation activity towards another histone tail. MOZ DPF binds unmodified H3 or H4 tails in chromatin and this interaction is stabilised by

acetylation of H3K9/14 by DPF+MYST. This recruitment could also involve the BRPF/ING5 complex proteins that could target MOZ within the vicinity of histone tails containing the H3K4me3 modification (Figure 6.2).

MOZ has also been found in complexes containing the methyltransferase MLL and the adaptor protein WDR5, which catalyses the trimethylation of H3K4 (Paggetti et al., 2010). Recruitment of MLL and WDR5 by MOZ may then drive the deposition of H3K4me3, which is incompatible with MOZ and thus, serves as a release signal from chromatin unless a free H4 tail is available. Conversely, MOZ could interact with chromatin bound MLL/WDR5, which facilitates MOZ acetylation of the same or neighbouring histone tails. Induced H4 acetylation serves as a second exit signal recognised by the MOZ DPF and so may then stimulate the dissociation of MOZ from chromatin (Figure 6.2).

Our studies implicate that MOZ functions within coactivator and corepressor complexes. Evidence now suggests that MOZ may also act to recruit other chromatin modulators to facilitate a repressive chromatin environment. Our data suggests this is through the recruitment of SUV39H1 by MOZ to the active coding region of genes, which consequently induces the trimethylation of H3K9, HP1 binding and the initiation of gene silencing (Figure 6.2). Thus, like several other chromatin-associated proteins, MOZ is a multi-functional regulator of chromatin modification and gene expression.



**Figure 6.2. Schematic model depicting the possible sequence of events involved in MOZ function.** MOZ coactivator complexes involve EAF6/BRPF/ING5 and MLL/WDR5 interactions. Evidence now suggests that MOZ also functions as a corepressor complex including SUV39H1 and HP1 interacting partners. Each complex consists of multiple ‘reader’ and ‘writer’ modules that must dictate the sequential order of PTM engagement and installation/removal, to signal for changes in gene expression.



## **Appendix**

Oligo Number	Oligo Name	Template	Restriction Enzyme	T <sub>m</sub> (°C)	Sequence (5' → 3')
LE240	MOZ 810-R	pcDNA3.1(+)-FLAG-MOZ	XbaI	56.4	AAAATCTAGATCACTCTAAATTCCTTTCCTG
LE576	GAL4-MOZ 509-F	pcDNA3.1(+)-FLAG-MOZ	BamHI	73.2	GGGCCCGGGGATCCCTCTGTGTCATTGAG
LE706	GST-TIP60 211-F	PET-TIP60	BamHI	80.2	AAAAGGATCCCGCATGACTGGCAGCCTG
LE707	GST-TIP60 513-R	PET-TIP60	EcoRI	77.1	AAAAGAAATTCGTGACCACCTCCCTCTTGTCTCC
LE708	GAL4-TIP60 211-F	PET-TIP60	BamHI	81.8	AAAAGGATCCGTGCGATGACTGGCAGCCTG
LE709	GAL4-TIP60 513-R	PET-TIP60	KpnI	78.3	AAAAGGTACCTGACCACCTCCCTCTTGTCTCC
LE710	GST-MOZ 509-F	pcDNA3.1(+)-FLAG-MOZ	BamHI	76.8	AAAAGGATCCCTCTGTGTCATTGAGTTGGG
LE711	GST-MOZ 810-R	pcDNA3.1(+)-FLAG-MOZ	EcoRI	69.7	AAAAGAAATTCGACCTCTAAATTCCTTTCCTG
LE712	GST-MORF 720-F	pMORF72	XhoI	72.8	AAATCGAGCCTTCGTGATTGAATTGG
LE713	GST-MORF 1001-R	pMORF72	BamHI	72.1	AAAAGGATCCCTGACTCTTCTTCAGACACTGC
LE714	GST-P300 1195-F	pGEX-p300 1097-1722	XhoI	67.8	AAAATCGAGGAGAAACAGGATCATCTTC
LE715	GST-P300 1290-F	pGEX-p300 1097-1722	XhoI	73.0	AAAATCGAGGCTAAAAGGTTGCCATCTACC
LE716	GST-P300 1660-R	pGEX-p300 1097-1722	EcoRI	78.7	AAAAGAAATTCGTGACGTGTGCAGCTCCACCAGC
LE717	GST-P300 1673-R	pGEX-p300 1097-1722	EcoRI	73.0	AAAAGAAATTCGTGAGCATTCATTGCAGGTGTAGAC
LE718	GAL4-P300 1195-F	pGEX-p300 1097-1722	KpnI	70.9	AAAAGGTACCCCTCAGAACAGGATCATCTTC
LE719	GAL4-P300 1290-F	pGEX-p300 1097-1722	KpnI	75.9	AAAAGGTACCCCTGCTAAAAGGTTGCCATCTACC
LE720	GAL4-P300 1660-R	pGEX-p300 1097-1722	SacI	80.6	AAAAGAGCTCTGACCGTGTGCAGCTCCACCAGC

LE721	GAL4-P300 1673-R	pGEX-p300 1097-1722	<u>SacI</u>	74.8	AAAGAGCTCTGAGCATTCATTGCAGGTGTAGAC
LE724	GST-CBP 1698-R	<u>pGEX-moz</u> -CBP 1098-1758	<u>BamHI</u>	80.9	AAAAGGAATCCGTGAGGTGTGAGCTCCACCAGC
LE725	GAL4-CBP 1232-F	<u>pGEX-moz</u> -CBP 1098-1758	<u>BamHI</u>	69.1	AAAAGGAATCCGTAATAGGATATCATTTCTGIGAG
LE726	GAL4-CBP 1318-F	<u>pGEX-moz</u> -CBP 1098-1758	<u>BamHI</u>	73.8	AAAAGGAATCCGTCCTCGAAAAAGAAAAAATTC
LE727	GAL4-CBP 1697-R	<u>pGEX-moz</u> -CBP 1098-1758	<u>KpnI</u>	79.3	AAAAGGTACCTGAGGTGTGAGCTCCACCAGC
LE732	GAL4-MORF 720-F	pMORF72	<u>BamHI</u>	75.0	AAAAGGAATCCGTCCTCTGTGATTGAATTGG
LE733	GAL4-MORF 1001-R	pMORF72	<u>KpnI</u>	70.4	AAAAGGTACCTGACTCTCTCTCAGACACTGC
LE734	GST-CBP 1232-F	<u>pGEX-moz</u> -CBP 1098-1758	<u>BglII</u>	63.5	AAAAGATCTCAGAAATAGGTATCATTTCTG
LE735	GST-CBP 1318-F	<u>pGEX-moz</u> -CBP 1098-1758	<u>BglII</u>	65.4	AAAAGATCTAGACCTCGGAAAGAAAC
LE738	GST-MOZ 776-R	pcDNA3.1(+)-FLAG-MOZ	<u>EcoRI</u>	75.5	AAAAGAAATCTGAOCAGCGCAACATTTCTGG
LE739	GAL4-MOZ 776-R	pcDNA3.1(+)-FLAG-MOZ	<u>XbaI</u>	74.0	AAAATCTAGATGAOCAGCGCAACATTTCTGG
LE787	MOZ 779-R	pcDNA3.1(+)-FLAG-MOZ	<u>XbaI</u>	74.9	AAAATCTAGATCAGACTGGAGTCCAGCGCAAAC
LE788	MOZ 770-R	pcDNA3.1(+)-FLAG-MOZ	<u>XbaI</u>	65.2	AAAATCTAGATCAATCTACATCTACAGGTGG
LE789	MOZ 784-R	pcDNA3.1(+)-FLAG-MOZ	<u>XbaI</u>	65.3	AAAATCTAGATCAAGAGTTGGACACTATGAC
LE884	FLAG-MOZ 510-F	pcDNA3.1(+)-FLAG-MOZ	<u>EcoRI</u>	88.1	AAAAGAAATCCACCATGGACTACAAGGACGACGAG ATAAGGGGTCTGTCAATTGAGTTTGGG
LE885	FLAG-MOZ 321-F	pcDNA3.1(+)-FLAG-MOZ	<u>EcoRI</u>	89.2	AAAAGAAATCCACCATGGACTACAAGGACGACGAG ATAAGGGGCAAAAGAGGACGACAG
LE891	FLAG-MOZ 1317-F	pcDNA3.1(+)-FLAG-MOZ	<u>EcoRI</u>	89.5	AAAAGAAATCCACCATGGACTACAAGGACGACGAG ATAAGGGGGATGATGAGGATGATGTC



LE892	MOZ 2004-R	pcDNA3.1(+)-FLAG-MOZ	<b>Xba</b> I	61.8	AAAATCTAGATCATCTTCTCATGTAAAG
LE893	FLAG-MOZ 810-F	pcDNA3.1(+)-FLAG-MOZ	<b>Eco</b> RI	88.5	AAAAGAAATTCACCATGGACTACAAGGACGACGAG ATAAGGGGAGATCAGTGTGGAAAG
LE894	MOZ 1117-R	pcDNA3.1(+)-FLAG-MOZ	<b>Xba</b> I	67.8	AAAATCTAGATCAATCAGCATCATCTGACTC
LE895	FLAG-MOZ 194-F	pcDNA3.1(+)-FLAG-MOZ	<b>Hind</b> III	87.9	AAAAAAGCTTCACCATGGACTACAAGGACGACGAG ATAAGGGGCTTCCACATGAAGG
LE896	MOZ 194-F	pcDNA3.1(+)-FLAG-MOZ	<b>Xho</b> I	77.1	AAACTCGAGCTTCCACATGAAAAGGATAAGCCG
LE897	MOZ 265-R	pcDNA3.1(+)-FLAG-MOZ	<b>Bam</b> HI	70.8	AAAAGGATCCTCAGCATGTTTACACTCG
LE898	MOZ 323-R	pcDNA3.1(+)-FLAG-MOZ	<b>Bam</b> HI	64.6	AAAAGGATCCTCACTTCTTTGTAGAAG
LE904	MOZ 510-R	pcDNA3.1(+)-FLAG-MOZ	<b>Xba</b> I	76.5	AAAATCTAGATCAAGAGGGACAGCGGACTTGTGG
LE905	MOZ 430-R	pcDNA3.1(+)-FLAG-MOZ	<b>Xba</b> I	79.0	AAAATCTAGATCACACTTCCCCCGAGCTTCCG
LE906	FLAG-MOZ 410-F	pcDNA3.1(+)-FLAG-MOZ	<b>Eco</b> RI	86.9	AAAAGAAATTCACCATGGACTACAAGGACGACGAG ATAAGGGGTTGAAGTTCAACAAGAAAACC
LE907	FLAG-MOZ 250-F	pcDNA3.1(+)-FLAG-MOZ	<b>Hind</b> III	89.8	AAAAAAGCTTCACCATGGACTACAAGGACGACGAG ATAAGGGGTTTCGAGTGAAGGCTTACCG
LE909	MOZ 250-F	pcDNA3.1(+)-FLAG-MOZ	<b>Xho</b> I	76.9	AAAACTCGAGGTTTCGAGTGAAGGCTTACCG

**Table A.1. Oligonucleotide Sequences Used for Cloning Purposes.** Oligonucleotides utilised during the research described in this thesis were obtained from Sigma Genosys and the sequences listed as detailed. Bold letters within the sequence indicate the restriction endonuclease site.

Oligo. Number	Oligo. Name	Template	T <sub>m</sub> (°C)	Sequence (5' → 3')
LE785	MOZ V779A/V780A-F	pcDNA3.1(+)-FLAG-MOZ	88.0	TTG CGCTGG ACT CCA <b>GCC GCA</b> GTGTCC AACTCT GTG
LE786	MOZ V779A/V781A-F	pcDNA3.1(+)-FLAG-MOZ	87.2	TTGCGCTGGACTCCA <b>GCC</b> ATA <b>GCG</b> GTCCA ACTCTGTGGTC
LE790	MOZ V779A/V780A - R	pcDNA3.1(+)-FLAG-MOZ	88.0	CACAGAGTTGGACAC <b>TCG GGC</b> TGGAGTCCAGCGCAA
LE791	MOZ V779A/V781A-R	pcDNA3.1(+)-FLAG-MOZ	87.2	GACCACAGAGTTGGA <b>GCG</b> TAT <b>GCG</b> TGGAGTCCAGCGCAA
LE827	MOZ C233A-F	pcDNA3.1(+)-FLAG-MOZ	91.5	GAACTCATCTCCTGTGCCGAC <b>GCT</b> GGCAACAAGTGGCCATCCATCC
LE828	MOZ C233A-R	pcDNA3.1(+)-FLAG-MOZ	91.5	GGATGGATGCCCACTGTTGCC <b>AGC</b> GTCCGCA CAGGAGATGAGTTC
LE829	MOZ C307A-F	pcDNA3.1(+)-FLAG-MOZ	78.9	CCAAAAGGCATGTGGATA <b>GCT</b> CAAAATATGTCGACCTAGG
LE830	MOZ C307A-R	pcDNA3.1(+)-FLAG-MOZ	78.9	CCTAGGTCGACATATTG <b>AGC</b> TATCCACATGCCCTTTTGG
LE840	MOZ P189A/V191A-F	pcDNA3.1(+)-FLAG-MOZ	82.0	GTCTCTTCCCTGTTT <b>AGCT</b> CCAG <b>GCG</b> TCCCTTCTTCCACATG
LE841	MOZ P189A/V191A-R	pcDNA3.1(+)-FLAG-MOZ	82.0	CATGTGGAAGAAGGA <b>GCG</b> TGTG <b>AGC</b> TAAACAGGAAAGAGAC
LE867	MOZ W257E-F	pcDNA3.1(+)-FLAG-MOZ	87.9	CGGTTGAGTGAAGGCCCTTACGG <b>GAG</b> CAGTGCATCGAGTGTAAACAT G
LE868	MOZ W257E-R	pcDNA3.1(+)-FLAG-MOZ	87.9	CATGTTTACACTCGATGCACTG <b>CT</b> CCCGTAAGGCCCTTCACCTCGAACCG
LE869	MOZ W305E-F	pcDNA3.1(+)-FLAG-MOZ	82.7	CCCGTATGCCAAAGGCATG <b>GAG</b> ATAATGTCAAATAATGTCGACC
LE870	MOZ W305E-R	pcDNA3.1(+)-FLAG-MOZ	82.7	GGTCGACATAATTGACATAT <b>CT</b> CCATGCCCTTTGGCATACGGG
LE899	MOZ 208-263 DELETION-F	pcDNA3.1(+)-FLAG-MOZ	82	GAGCTGTCTCGACGTACAAAACCCCTAACCAAGTCGTTGGCOG
LE900	MOZ 208-263 DELETION-R	pcDNA3.1(+)-FLAG-MOZ	86.7	CGGCCAACGACTTGTTAGGGGTTTGTACGTCGAGGACAGCTC

**Table A.2. Oligonucleotide Sequences Used for Site-Directed Mutagenesis.** Oligonucleotides utilised during

the research described in this thesis were obtained from Sigma Genosys and the sequences listed as detailed. Bold letters within the sequence indicate the desired amino acid codon change within the sequence.

<b>Vector</b>	<b>Insert</b>	<b>Source</b>
pcDNA3.0-PT	FLAG-MOZ (full-length)	Phillip Troke
pcDNA3.0-PT	FLAG-MOZ N	Phillip Troke
pGEX-DMH	MOZ 1-321	Phillip Troke
pGEX-DMH	MOZ 1-171	Phillip Troke
pGEX-DMH	MOZ 172-321	Phillip Troke
pEYFP-C1	YFP-FLAG-MOZ	Phillip Troke
pcDNA3.1	HA-MOZ 1-171	Nijjareeya Sirisriro
pcDNA3.1	FLAG-MOZ 1-321	Nijjareeya Sirisriro
pcDNA3.1	FLAG-MOZ-TIF2	Phillip Troke
pSG5	mHP1 $\alpha$	B. Le Douarin
pSG5	mHP1 $\beta$	B. Le Douarin
pCMV	(Myc)3-SUV39H1	T. Jenuwein
pCMV	(Myc)3-SUV39H1 H324L	T. Jenuwein
pGEX	SUV39H1	T. Jenuwein
pGEX	SUV39H1 H324L	T. Jenuwein
pcDNA3.1(-)	FLAG-MORF	Xiang-Jiao Yang
pSG424 (GAL4)	-	Martin Dickens
PET3a	TIP60	Eric Kalkhoven
pGEX	CBP 1098-1758	Andrew Bannister
pGEX	p300 1097-1722	Andrew Bannister
pCH110	Lac-Z	Pharmacia
pT109	3xAML1-luc	Arthur Zelent
pCMV5	AML1	D.E Zhong
peYFP (C1)	HP1 $\gamma$	Angus Lamond
pSG5	TDG	Simak Ali
pcDNA3.1	FLAG-HIPK1	Hilary Collins
pGEX-2T	SUMO-1	Ron Hay
pGEX-2T	SUMO-2	Ron Hay

**Table A.3. Origin of Plasmids.** Plasmids utilised during the research described in this thesis were obtained as detailed.



<b>Insert</b>	<b>Vector</b>	<b>Mutation</b>
GAL4-MOZ HAT	pSG424	×
GAL4-MORF HAT	pSG424	×
GAL4-P300 HAT	pSG424	×
GST-TIP60 HAT	pGEX-DMH	×
GAL4-TIP60 HAT	pSG424	×
GST-MOZ HAT	pGEX-DMH	×
GST-MOZ 510-776	pGEX-DMH	×
GST-MORF HAT	pGEX-DMH	×
GST-PCAF HAT	pGEX-DMH	×
GAL4-PCAF HAT	pSG424	×
FLAG-MOZ 321-770	pcDNA3.1 +	×
FLAG-MOZ 321-810	pcDNA3.1 +	×
FLAG-MOZ 510-770	pcDNA3.1 +	×
FLAG-MOZ 510-810	pcDNA3.1 +	×
FLAG-MOZ 810-1117	pcDNA3.1 +	×
FLAG-MOZ 1317-2004	pcDNA3.1 +	×
FLAG-MOZ 1-770	pcDNA3.1 +	×
FLAG-MOZ 1-779	pcDNA3.1 +	×
FLAG-MOZ 1-784	pcDNA3.1 +	×
FLAG-MOZ 1-810	pcDNA3.1 +	×
FLAG-MOZ 321-430	pcDNA3.1 +	×
FLAG-MOZ 410-510	pcDNA3.1 +	×
FLAG-MOZ 194-265	pcDNA3.1 +	×
FLAG-MOZ 194-323	pcDNA3.1 +	×
FLAG-MOZ N (1-1117)	pcDNA3.1 +	V779A/I780A
FLAG-MOZ N (1-1117)	pcDNA3.1 +	V779A/V781A
GST-MOZ 194-810	pGEX-DMH	×
GST-MOZ 510-810	pGEX-DMH	V779A/V781A
GST-MOZ 194-265	pGEX-DMH	×

GST-MOZ 194-323	pGEX-DMH	✕
GST-MOZ 250-323	pGEX-DMH	✕
FLAG-MOZ 321-510	pcDNA3.1 +	✕
FLAG-MOZ N (1-1117)	pcDNA3.1 +	C233A
FLAG-MOZ N (1-1117)	pcDNA3.1 +	C307A
FLAG-MOZ N (1-1117)	pcDNA3.1 +	C233A/C307A
FLAG-MOZ N (1-1117)	pcDNA3.1 +	W257E
FLAG-MOZ N (1-1117)	pcDNA3.1 +	W305E
FLAG-MOZ N (1-1117)	pcDNA3.1 +	W257E/W305E
FLAG-MOZ N (1-1117)	pcDNA3.1 +	P189A/V191A
FLAG-MOZ	pcDNA3.1 +	C233A
FLAG-MOZ	pcDNA3.1 +	C307A
FLAG-MOZ	pcDNA3.1 +	W257E
FLAG-MOZ-TIF2	pcDNA3.1 +	C233A
FLAG-MOZ-TIF2	pcDNA3.1 +	W257E
FLAG-MOZ-TIF2	pcDNA3.1 +	W305E
FLAG-MOZ-TIF2	pcDNA3.1 +	W257E/W305E
FLAG-MOZ-TIF2	pcDNA3.1 +	P189A/V191A
GST-MOZ 1-321	pGEX-DMH	C233A
GST-MOZ 1-321	pGEX-DMH	C307A
GST-MOZ 1-321	pGEX-DMH	C233A/C307A
GST-MOZ 1-321	pGEX-DMH	W257E
GST-MOZ 1-321	pGEX-DMH	W305E
GST-MOZ 1-321	pGEX-DMH	W257E/W305E
GST-MOZ 1-321	pGEX-DMH	P189A/V191A
GST-MOZ 1-172	pGEX-DMH	P189A/V191A
GST-MORF 201-330	pGEX-DMH	✕
GST-MORF 201-1001	pGEX-DMH	✕

**Table A.4. Plasmids Created by Cloning or Site-directed Mutagenesis.** Plasmids created and utilised during the research described in this thesis are listed above.



	<b>Antibody</b>	<b>Dilution</b>		<b>Antibody</b>	<b>Dilution</b>
<b>1</b>	H2A Abcam	1:250	<b>14</b>	H4K16ac Upstate	1:500
<b>2</b>	H2B Abcam	1:250	<b>15</b>	H4 acetyl-lysine Millipore	1:500WB
<b>3</b>	H3 Abcam	1:1000	<b>16</b>	GST Sigma	1:1000
<b>4</b>	H3K4me3 Abcam	1:150 IF 1:500 WB	<b>17</b>	MOZ (N-19) Santa Cruz	1:150 IF 1:500 WB
<b>5</b>	H3K9ac Abcam	1:150 IF 1:250 WB	<b>18</b>	MOZ 127A LLR	1:500 WB 2 µg IP
<b>6</b>	H3K9me3 Abcam	1:150 IF 1:500 WB	<b>19</b>	FLAG Sigma	1:500 IF 1:1000 WB
<b>7</b>	H3K14ac Upstate	1:150 IF 1:500 WB	<b>20</b>	FLAG Sigma	2 µg chIP/ IP
<b>8</b>	H3 acetyl-lysine Millipore	1:150 IF 1:500 WB	<b>21</b>	c-Myc Santa Cruz	1:150 IF 1:250 WB
<b>9</b>	H4 Abcam	1:250	<b>22</b>	SUV39H1 Upstate	1:500
<b>10</b>	H4 Millipore	1:500	<b>23</b>	AML1 CalBiochem	1:200
<b>11</b>	H4K5ac Upstate	1:500	<b>24</b>	GFP Abcam	1:1000
<b>12</b>	H4K8ac Upstate	1:500	<b>25</b>	SUMO 1 Santa Cruz	1:500
<b>13</b>	H4K12ac Upstate	1:500	<b>26</b>	SUMO 2/3 Santa Cruz	1:500

**Table A.5. Origin and Dilutions of Antibodies.** Antibodies utilised during the research described in this thesis were obtained as detailed.

## References

- Aguilera, A. 2002. The connection between transcription and genomic instability. *EMBO J.* 21:195-201.
- Aikawa, Y., T. Katsumoto, P. Zhang, H. Shima, M. Shino, K. Terui, E. Ito, H. Ohno, E.R. Stanley, H. Singh, D.G. Tenen, and I. Kitabayashi. 2010. PU.1-mediated upregulation of CSF1R is crucial for leukemia stem cell potential induced by MOZ-TIF2. *Nat Med.* 16:580-5, 1p following 585.
- Allis, C.D., S.L. Berger, J. Cote, S. Dent, T. Jenuwien, T. Kouzarides, L. Pillus, D. Reinberg, Y. Shi, R. Shiekhata, A. Shilatifard, J. Workman, and Y. Zhang. 2007. New nomenclature for chromatin-modifying enzymes. *Cell.* 131:633-6.
- Anamika, K., A.R. Krebs, J. Thompson, O. Poch, D. Devys, and L. Tora. 2010. Lessons from genome-wide studies: an integrated definition of the coactivator function of histone acetyl transferases. *Epigenetics Chromatin.* 3:18.
- Argentaro, A., J.C. Yang, L. Chapman, M.S. Kowalczyk, R.J. Gibbons, D.R. Higgs, D. Neuhaus, and D. Rhodes. 2007. Structural consequences of disease-causing mutations in the ATRX-DNMT3-DNMT3L (ADD) domain of the chromatin-associated protein ATRX. *Proc Natl Acad Sci U S A.* 104:11939-44.
- Arif, M., G.V. Kumar, C. Narayana, and T.K. Kundu. 2007. Autoacetylation induced specific structural changes in histone acetyltransferase domain of p300: probed by surface enhanced Raman spectroscopy. *J Phys Chem B.* 111:11877-9.
- Arif, M., S.K. Pradhan, G.R. Thanuja, B.M. Vedomurthy, S. Agrawal, D. Dasgupta, and T.K. Kundu. 2009. Mechanism of p300 specific histone acetyltransferase inhibition by small molecules. *J Med Chem.* 52:267-77.
- Avvakumov, N., and J. Cote. 2007. The MYST family of histone acetyltransferases and their intimate links to cancer. *Oncogene.* 26:5395-407.
- Ayoub, N., A.D. Jeyasekharan, J.A. Bernal, and A.R. Venkitaraman. 2008. HP1-beta mobilization promotes chromatin changes that initiate the DNA damage response. *Nature.* 453:682-6.
- Ayton, P.M., and M.L. Cleary. 2003. Transformation of myeloid progenitors by MLL oncoproteins is dependent on Hoxa7 and Hoxa9. *Genes Dev.* 17:2298-307.
- Badugu, R., M.M. Shareef, and R. Kellum. 2003. Novel Drosophila heterochromatin protein 1 (HP1)/origin recognition complex-associated protein (HOAP) repeat motif in HP1/HOAP interactions and chromocenter associations. *J Biol Chem.* 278:34491-8.
- Baker, L.A., C.D. Allis, and G.G. Wang. 2008. PHD fingers in human diseases: disorders arising from misinterpreting epigenetic marks. *Mutat Res.* 647:3-12.
- Balasubramanyam, K., V. Swaminathan, A. Ranganathan, and T.K. Kundu. 2003. Small molecule modulators of histone acetyltransferase p300. *J Biol Chem.* 278:19134-40.
- Balasubramanyam, K., R.A. Varier, M. Altaf, V. Swaminathan, N.B. Siddappa, U. Ranga, and T.K. Kundu. 2004. Curcumin, a novel p300/CREB-binding protein-specific inhibitor of acetyltransferase, represses the acetylation of histone/nonhistone proteins and histone acetyltransferase-dependent chromatin transcription. *J Biol Chem.* 279:51163-71.
- Bannister, A.J., and T. Kouzarides. 2005. Reversing histone methylation. *Nature.* 436:1103-6.

- Bartova, E., J. Krejci, A. Harnicarova, G. Galiova, and S. Kozubek. 2008a. Histone Modifications and Nuclear Architecture: A Review. *J Histochem Cytochem*.
- Bartova, E., J. Krejci, A. Harnicarova, and S. Kozubek. 2008b. Differentiation of human embryonic stem cells induces condensation of chromosome territories and formation of heterochromatin protein 1 foci. *Differentiation*. 76:24-32.
- Bedford, M.T., and S. Richard. 2005. Arginine methylation an emerging regulator of protein function. *Mol Cell*. 18:263-72.
- Berndsen, C.E., B.N. Albaugh, S. Tan, and J.M. Denu. 2007. Catalytic mechanism of a MYST family histone acetyltransferase. *Biochemistry*. 46:623-9.
- Bienz, M. 2006. The PHD finger, a nuclear protein-interaction domain. *Trends Biochem Sci*. 31:35-40.
- Blanco, J.C., S. Minucci, J. Lu, X.J. Yang, K.K. Walker, H. Chen, R.M. Evans, Y. Nakatani, and K. Ozato. 1998. The histone acetylase PCAF is a nuclear receptor coactivator. *Genes Dev*. 12:1638-51.
- Borrow, J., V.P. Stanton, Jr., J.M. Andresen, R. Becher, F.G. Behm, R.S. Chaganti, C.I. Civin, C. Disteche, I. Dube, A.M. Frischauf, D. Horsman, F. Mitelman, S. Volinia, A.E. Watmore, and D.E. Housman. 1996. The translocation t(8;16)(p11;p13) of acute myeloid leukaemia fuses a putative acetyltransferase to the CREB-binding protein. *Nat Genet*. 14:33-41.
- Braig, M., S. Lee, C. Loddenkemper, C. Rudolph, A.H. Peters, B. Schlegelberger, H. Stein, B. Dorken, T. Jenuwein, and C.A. Schmitt. 2005. Oncogene-induced senescence as an initial barrier in lymphoma development. *Nature*. 436:660-5.
- Bryan, E.J., V.J. Jokubaitis, N.L. Chamberlain, S.W. Baxter, E. Dawson, D.Y. Choong, and I.G. Campbell. 2002. Mutation analysis of EP300 in colon, breast and ovarian carcinomas. *Int J Cancer*. 102:137-41.
- Buczek-Thomas, J.A., E. Hsia, C.B. Rich, J.A. Foster, and M.A. Nugent. 2008. Inhibition of histone acetyltransferase by glycosaminoglycans. *J Cell Biochem*.
- Cairns, B.R. 2009. The logic of chromatin architecture and remodelling at promoters. *Nature*. 461:193-8.
- Camos, M., J. Esteve, P. Jares, D. Colomer, M. Rozman, N. Villamor, D. Costa, A. Carrio, J. Nomdedeu, E. Montserrat, and E. Campo. 2006. Gene expression profiling of acute myeloid leukemia with translocation t(8;16)(p11;p13) and MYST3-CREBBP rearrangement reveals a distinctive signature with a specific pattern of HOX gene expression. *Cancer Res*. 66:6947-54.
- Canella, D., V. Praz, J.H. Reina, P. Cousin, and N. Hernandez. 2010. Defining the RNA polymerase III transcriptome: Genome-wide localization of the RNA polymerase III transcription machinery in human cells. *Genome Res*. 20:710-21.
- Canzio, D., E.Y. Chang, S. Shankar, K.M. Kuchenbecker, M.D. Simon, H.D. Madhani, G.J. Narlikar, and B. Al-Sady. 2011. Chromodomain-mediated oligomerization of HP1 suggests a nucleosome-bridging mechanism for heterochromatin assembly. *Mol Cell*. 41:67-81.
- Carapeti, M., R.C. Aguiar, J.M. Goldman, and N.C. Cross. 1998. A novel fusion between MOZ and the nuclear receptor coactivator TIF2 in acute myeloid leukemia. *Blood*. 91:3127-33.

- Carter, D.R., C. Eskiw, and P.R. Cook. 2008. Transcription factories. *Biochem Soc Trans.* 36:585-9.
- Chakraborty, S., K.K. Sinha, V. Senyuk, and G. Nucifora. 2003. SUV39H1 interacts with AML1 and abrogates AML1 transactivity. AML1 is methylated in vivo. *Oncogene.* 22:5229-37.
- Champagne, K.S., N. Saksouk, P.V. Pena, K. Johnson, M. Ullah, X.J. Yang, J. Cote, and T.G. Kutateladze. 2008. The crystal structure of the ING5 PHD finger in complex with an H3K4me3 histone peptide. *Proteins.* 72:1371-6.
- Champagne, N., N. Pelletier, and X.J. Yang. 2001. The monocytic leukemia zinc finger protein MOZ is a histone acetyltransferase. *Oncogene.* 20:404-9.
- Chan, E.M., R.J. Chan, E.M. Comer, R.J. Goulet, 3rd, C.D. Crean, Z.D. Brown, A.M. Fruehwald, Z. Yang, H.S. Boswell, H. Nakshatri, and T.G. Gabig. 2007. MOZ and MOZ-CBP cooperate with NF-kappaB to activate transcription from NF-kappaB-dependent promoters. *Exp Hematol.* 35:1782-92.
- Chen, J., F.M. Ghazawi, and Q. Li. 2010. Interplay of bromodomain and histone acetylation in the regulation of p300-dependent genes. *Epigenetics.* 5:509-15.
- Chen, J., D.A. Santillan, M. Koonce, W. Wei, R. Luo, M.J. Thirman, N.J. Zeleznik-Le, and M.O. Diaz. 2008. Loss of MLL PHD finger 3 is necessary for MLL-ENL-induced hematopoietic stem cell immortalization. *Cancer Res.* 68:6199-207.
- Chen, R., V.H. Kang, J. Chen, J.C. Shope, J. Torabinejad, D.B. DeWald, and G.D. Prestwich. 2002. A monoclonal antibody to visualize PtdIns(3,4,5)P(3) in cells. *J Histochem Cytochem.* 50:697-708.
- Cherrier, T., S. Suzanne, L. Redel, M. Calao, C. Marban, B. Samah, R. Mukerjee, C. Schwartz, G. Gras, B.E. Sawaya, S.L. Zeichner, D. Aunis, C. Van Lint, and O. Rohr. 2009. p21(WAF1) gene promoter is epigenetically silenced by CTIP2 and SUV39H1. *Oncogene.* 28:3380-9.
- Chignola, F., M. Gaetani, A. Rebane, T. Org, L. Mollica, C. Zucchelli, A. Spitaleri, V. Mannella, P. Peterson, and G. Musco. 2009. The solution structure of the first PHD finger of autoimmune regulator in complex with non-modified histone H3 tail reveals the antagonistic role of H3R2 methylation. *Nucleic Acids Res.* 37:2951-61.
- Choudhary, C., C. Kumar, F. Gnad, M.L. Nielsen, M. Rehman, T.C. Walther, J.V. Olsen, and M. Mann. 2009. Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science.* 325:834-40.
- Chrivia, J.C., R.P. Kwok, N. Lamb, M. Hagiwara, M.R. Montminy, and R.H. Goodman. 1993. Phosphorylated CREB binds specifically to the nuclear protein CBP. *Nature.* 365:855-9.
- Churchman, L.S., J.S. Weissman. 2011. Nascent transcript sequencing visualizes transcription at nucleotide resolution. *Nature.* 469:368-73.
- Clapier, C.R., and B.R. Cairns. 2009. The biology of chromatin remodeling complexes. *Annu Rev Biochem.* 78:273-304.
- Clarkson, M.J., J.R. Wells, F. Gibson, R. Saint, and D.J. Tremethick. 1999. Regions of variant histone His2AvD required for Drosophila development. *Nature.* 399:694-7.
- Collins, H.M., K.B. Kindle, S. Matsuda, C. Ryan, P.J. Troke, E. Kalkhoven, and D.M. Heery. 2006. MOZ-TIF2 alters cofactor recruitment and histone

- modification at the RARbeta2 promoter: differential effects of MOZ fusion proteins on CBP- and MOZ-dependent activators. *J Biol Chem.* 281:17124-33.
- Cook, P.R. 1999. The organization of replication and transcription. *Science.* 284:1790-5.
- Core, L.J., and J.T. Lis. 2008. Transcription regulation through promoter-proximal pausing of RNA polymerase II. *Science.* 319:1791-2.
- Core, L.J., Waterfall, J.J., and J.T. Lis. 2008. Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science.* 322:1845-8.
- Cowieson, N.P., J.F. Partridge, R.C. Allshire, and P.J. McLaughlin. 2000. Dimerisation of a chromo shadow domain and distinctions from the chromodomain as revealed by structural analysis. *Curr Biol.* 10:517-25.
- Crump, J.G., M.E. Swartz, J.K. Eberhart, and C.B. Kimmel. 2006. Moz-dependent Hox expression controls segment-specific fate maps of skeletal precursors in the face. *Development.* 133:2661-9.
- Cuthbert, G.L., S. Daujat, A.W. Snowden, H. Erdjument-Bromage, T. Hagiwara, M. Yamada, R. Schneider, P.D. Gregory, P. Tempst, A.J. Bannister, and T. Kouzarides. 2004. Histone deimination antagonizes arginine methylation. *Cell.* 118:545-53.
- Daser, A., and T.H. Rabbitts. 2005. The versatile mixed lineage leukaemia gene MLL and its many associations in leukaemogenesis. *Semin Cancer Biol.* 15:175-88.
- Dawson, M.A., A.J. Bannister, B. Gottgens, S.D. Foster, T. Bartke, A.R. Green, and T. Kouzarides. 2009. JAK2 phosphorylates histone H3Y41 and excludes HP1alpha from chromatin. *Nature.* 461:819-22.
- de Wit, E., F. Greil, and B. van Steensel. 2007. High-resolution mapping reveals links of HP1 with active and inactive chromatin components. *PLoS Genet.* 3:e38.
- Deguchi, K., P.M. Ayton, M. Carapeti, J.L. Kutok, C.S. Snyder, I.R. Williams, N.C. Cross, C.K. Glass, M.L. Cleary, and D.G. Gilliland. 2003. MOZ-TIF2-induced acute myeloid leukemia requires the MOZ nucleosome binding motif and TIF2-mediated recruitment of CBP. *Cancer Cell.* 3:259-71.
- Dekker, F.J., and H.J. Haisma. 2009. Histone acetyl transferases as emerging drug targets. *Drug Discov Today.* 14:942-8.
- Dilworth, F.J., K.J. Seaver, A.L. Fishburn, S.L. Htet, and S.J. Tapscott. 2004. In vitro transcription system delineates the distinct roles of the coactivators pCAF and p300 during MyoD/E47-dependent transactivation. *Proc Natl Acad Sci U S A.* 101:11593-8.
- Dinant, C., and M.S. Luijsterburg. 2009. The emerging role of HP1 in the DNA damage response. *Mol Cell Biol.* 29:6335-40.
- Dorigo, B., Schalch, T., Kulangara, A., Duda, S., Schroeder, R.R., T.J. Richmond. 2004. Nucleosome arrays reveal the two-start organization of the chromatin fiber. *Science.* 306:1571-3.
- Doyon, Y., C. Cayrou, M. Ullah, A.J. Landry, V. Cote, W. Selleck, W.S. Lane, S. Tan, X.J. Yang, and J. Cote. 2006. ING tumor suppressor proteins are critical regulators of chromatin acetylation required for genome expression and perpetuation. *Mol Cell.* 21:51-64.

- Drygin, D., W.G. Rice, and I. Grummt. 2010. The RNA polymerase I transcription machinery: an emerging target for the treatment of cancer. *Annu Rev Pharmacol Toxicol.* 50:131-56.
- Eckner, R., M.E. Ewen, D. Newsome, M. Gerdes, J.A. DeCaprio, J.B. Lawrence, and D.M. Livingston. 1994. Molecular cloning and functional analysis of the adenovirus E1A-associated 300-kD protein (p300) reveals a protein with properties of a transcriptional adaptor. *Genes Dev.* 8:869-84.
- Eissenberg, J.C., T.C. James, D.M. Foster-Hartnett, T. Hartnett, V. Ngan, and S.C. Elgin. 1990. Mutation in a heterochromatin-specific chromosomal protein is associated with suppression of position-effect variegation in *Drosophila melanogaster*. *Proc Natl Acad Sci U S A.* 87:9923-7.
- Eitoku, M., L. Sato, T. Senda, and M. Horikoshi. 2008. Histone chaperones: 30 years from isolation to elucidation of the mechanisms of nucleosome assembly and disassembly. *Cell Mol Life Sci.* 65:414-44.
- Epstein, D.J. 2009. Cis-regulatory mutations in human disease. *Brief Funct Genomic Proteomic.* 8:310-6.
- Esteyries, S., C. Perot, J. Adelaide, M. Imbert, A. Lagarde, C. Pautas, S. Olschwang, D. Birnbaum, M. Chaffanet, and M.J. Mozziconacci. 2008. NCOA3, a new fusion partner for MOZ/MYST3 in M5 acute myeloid leukemia. *Leukemia.* 22:663-5.
- Fanti, L., M. Berloco, L. Piacentini, and S. Pimpinelli. 2003. Chromosomal distribution of heterochromatin protein 1 (HP1) in *Drosophila*: a cytological map of euchromatic HP1 binding sites. *Genetica.* 117:135-47.
- Fanti, L., and S. Pimpinelli. 2008. HP1: a functionally multifaceted protein. *Curr Opin Genet Dev.* 18:169-74.
- Faro-Trindade, I., and P.R. Cook. 2006. Transcription factories: structures conserved during differentiation and evolution. *Biochem Soc Trans.* 34:1133-7.
- Ferreira, H., J. Somers, R. Webster, A. Flaus, and T. Owen-Hughes. 2007. Histone tails and the H3 alphaN helix regulate nucleosome mobility and stability. *Mol Cell Biol.* 27:4037-48.
- Fischle, W., B.S. Tseng, H.L. Dormann, B.M. Ueberheide, B.A. Garcia, J. Shabanowitz, D.F. Hunt, H. Funabiki, and C.D. Allis. 2005. Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation. *Nature.* 438:1116-22.
- Fraga, M.F., E. Ballestar, A. Villar-Garea, M. Boix-Chornet, J. Espada, G. Schotta, T. Bonaldi, C. Haydon, S. Ropero, K. Petrie, N.G. Iyer, A. Perez-Rosado, E. Calvo, J.A. Lopez, A. Cano, M.J. Calasanz, D. Colomer, M.A. Piris, N. Ahn, A. Imhof, C. Caldas, T. Jenuwein, and M. Esteller. 2005. Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. *Nat Genet.* 37:391-400.
- Frisch, S.M., and J.S. Mymryk. 2002. Adenovirus-5 E1A: paradox and paradigm. *Nat Rev Mol Cell Biol.* 3:441-52.
- Fuda, N.J., M.B. Ardehali, and J.T. Lis. 2009. Defining mechanisms that regulate RNA polymerase II transcription in vivo. *Nature.* 461:186-92.
- Fussner, E., R.W. Ching, and D.P. Bazett-Jones. 2010. Living without 30nm chromatin fibers. *Trends Biochem Sci.*

- Gabig, T.G., P.L. Mantel, R. Rosli, and C.D. Crean. 1994. Requiem: a novel zinc finger gene essential for apoptosis in myeloid cells. *J Biol Chem.* 269:29515-9.
- Gayther, S.A., S.J. Batley, L. Linger, A. Bannister, K. Thorpe, S.F. Chin, Y. Daigo, P. Russell, A. Wilson, H.M. Sowter, J.D. Delhanty, B.A. Ponder, T. Kouzarides, and C. Caldas. 2000. Mutations truncating the EP300 acetylase in human cancers. *Nat Genet.* 24:300-3.
- Gozani, O., P. Karuman, D.R. Jones, D. Ivanov, J. Cha, A.A. Lugovskoy, C.L. Baird, H. Zhu, S.J. Field, S.L. Lessnick, J. Villasenor, B. Mehrotra, J. Chen, V.R. Rao, J.S. Brugge, C.G. Ferguson, B. Payastre, D.G. Myszka, L.C. Cantley, G. Wagner, N. Divecha, G.D. Prestwich, and J. Yuan. 2003. The PHD finger of the chromatin-associated protein ING2 functions as a nuclear phosphoinositide receptor. *Cell.* 114:99-111.
- Grossman, S.R., M. Perez, A.L. Kung, M. Joseph, C. Mansur, Z.X. Xiao, S. Kumar, P.M. Howley, and D.M. Livingston. 1998. p300/MDM2 complexes participate in MDM2-mediated p53 degradation. *Mol Cell.* 2:405-15.
- Grummt, I. 1999. Regulation of mammalian ribosomal gene transcription by RNA polymerase I. *Prog Nucleic Acid Res Mol Biol.* 62:109-54.
- Guelman, S., T. Suganuma, L. Florens, S.K. Swanson, C.L. Kiesecker, T. Kusch, S. Anderson, J.R. Yates, 3rd, M.P. Washburn, S.M. Abmayr, and J.L. Workman. 2006. Host cell factor and an uncharacterized SANT domain protein are stable components of ATAC, a novel dAda2A/dGcn5-containing histone acetyltransferase complex in *Drosophila*. *Mol Cell Biol.* 26:871-82.
- Hahn, S. 2004. Structure and Mechanism of RNA Polymerase II Transcription Machinery. *Nature Structural & Molecular Biology.* 11:394-403.
- Hassa, P.O., S.S. Haenni, M. Elser, and M.O. Hottiger. 2006. Nuclear ADP-ribosylation reactions in mammalian cells: where are we today and where are we going? *Microbiol Mol Biol Rev.* 70:789-829.
- Hassan, Y.I., and J. Zemleni. 2008. A novel, enigmatic histone modification: biotinylation of histones by holocarboxylase synthetase. *Nutr Rev.* 66:721-5.
- Healy, S., B. Perez-Cadahia, D. Jia, M.K. McDonald, J.R. Davie, and R.A. Gravel. 2009. Biotin is not a natural histone modification. *Biochim Biophys Acta.* 1789:719-33.
- Heery, D.M., and P.M. Fischer. 2007. Pharmacological targeting of lysine acetyltransferases in human disease: a progress report. *Drug Discov Today.* 12:88-99.
- Herr, A.J., M.B. Jensen, T. Dalmay, and D.C. Baulcombe. 2005. RNA polymerase IV directs silencing of endogenous DNA. *Science.* 308:118-20.
- Holbert, M.A., T. Sikorski, J. Carten, D. Snowflack, S. Hodawadekar, and R. Marmorstein. 2007. The human monocytic leukemia zinc finger histone acetyltransferase domain contains DNA-binding activity implicated in chromatin targeting. *J Biol Chem.* 282:36603-13.
- Huang, Y., J. Fang, M.T. Bedford, Y. Zhang, and R.M. Xu. 2006. Recognition of histone H3 lysine-4 methylation by the double tudor domain of JMJD2A. *Science.* 312:748-51.
- Huntly, B.J., H. Shigematsu, K. Deguchi, B.H. Lee, S. Mizuno, N. Duclos, R. Rowan, S. Amaral, D. Curley, I.R. Williams, K. Akashi, and D.G. Gilliland. 2004.



- MOZ-TIF2, but not BCR-ABL, confers properties of leukemic stem cells to committed murine hematopoietic progenitors. *Cancer Cell*. 6:587-96.
- Iberg, A.N., A. Espejo, D. Cheng, D. Kim, J. Michaud-Levesque, S. Richard, and M.T. Bedford. 2008. Arginine methylation of the histone H3 tail impedes effector binding. *J Biol Chem*. 283:3006-10.
- Illingworth, R.S., and A.P. Bird. 2009. CpG islands--'a rough guide'. *FEBS Lett*. 583:1713-20.
- Imhof, A., X.J. Yang, V.V. Ogryzko, Y. Nakatani, A.P. Wolffe, and H. Ge. 1997. Acetylation of general transcription factors by histone acetyltransferases. *Curr Biol*. 7:689-92.
- Ivanov, A.V., H. Peng, V. Yurchenko, K.L. Yap, D.G. Negorev, D.C. Schultz, E. Psulkowski, W.J. Fredericks, D.E. White, G.G. Maul, M.J. Sadofsky, M.M. Zhou, and F.J. Rauscher, 3rd. 2007. PHD domain-mediated E3 ligase activity directs intramolecular sumoylation of an adjacent bromodomain required for gene silencing. *Mol Cell*. 28:823-37.
- Jacobson, R.H., A.G. Ladurner, D.S. King, and R. Tjian. 2000. Structure and function of a human TAFII250 double bromodomain module. *Science*. 288:1422-5.
- Joazeiro, C.A., and A.M. Weissman. 2000. RING finger proteins: mediators of ubiquitin ligase activity. *Cell*. 102:549-52.
- Jones, R.S., and W.M. Gelbart. 1993. The Drosophila Polycomb-group gene Enhancer of zeste contains a region with sequence similarity to trithorax. *Mol Cell Biol*. 13:6357-66.
- Ju, B.G., V.V. Lunyak, V. Perissi, I. Garcia-Bassets, D.W. Rose, C.K. Glass, and M.G. Rosenfeld. 2006. A topoisomerase II $\beta$ -mediated dsDNA break required for regulated transcription. *Science*. 312:1798-802.
- Kalkhoven, E. 2004. CBP and p300: HATs for different occasions. *Biochem Pharmacol*. 68:1145-55.
- Kalkhoven, E., J.H. Roelfsema, H. Teunissen, A. den Boer, Y. Ariyurek, A. Zantema, M.H. Breuning, R.C. Hennekam, and D.J. Peters. 2003. Loss of CBP acetyltransferase activity by PHD finger mutations in Rubinstein-Taybi syndrome. *Hum Mol Genet*. 12:441-50.
- Katsumoto, T., Y. Aikawa, A. Iwama, S. Ueda, H. Ichikawa, T. Ochiya, and I. Kitabayashi. 2006. MOZ is essential for maintenance of hematopoietic stem cells. *Genes Dev*. 20:1321-30.
- Katsumoto, T., N. Yoshida, and I. Kitabayashi. 2008. Roles of the histone acetyltransferase monocytic leukemia zinc finger protein in normal and malignant hematopoiesis. *Cancer Sci*. 99:1523-7.
- Kindle, K.B., H.M. Collins, and D.M. Heery. 2010. MOZ-TIF2-mediated destruction of CBP/p300 is blocked by calpain inhibitor 2. *Leukemia*. 24:1359-61.
- Kindle, K.B., P.J. Troke, H.M. Collins, S. Matsuda, D. Bossi, C. Bellodi, E. Kalkhoven, P. Salomoni, P.G. Pelicci, S. Minucci, and D.M. Heery. 2005. MOZ-TIF2 inhibits transcription by nuclear receptors and p53 by impairment of CBP function. *Mol Cell Biol*. 25:988-1002.
- Kitabayashi, I., Y. Aikawa, L.A. Nguyen, A. Yokoyama, and M. Ohki. 2001a. Activation of AML1-mediated transcription by MOZ and inhibition by the MOZ-CBP fusion protein. *EMBO J*. 20:7184-96.

- Kitabayashi, I., Y. Aikawa, A. Yokoyama, F. Hosoda, M. Nagai, N. Kakazu, T. Abe, and M. Ohki. 2001b. Fusion of MOZ and p300 histone acetyltransferases in acute monocytic leukemia with a t(8;22)(p11;q13) chromosome translocation. *Leukemia*. 15:89-94.
- Kobza, K., G. Camporeale, B. Rueckert, A. Kueh, J.B. Griffin, G. Sarath, and J. Zemleni. 2005. K4, K9 and K18 in human histone H3 are targets for biotinylation by biotinidase. *FEBS J*. 272:4249-59.
- Komander, D. 2009. The emerging complexity of protein ubiquitination. *Biochem Soc Trans*. 37:937-53.
- Kornberg, R.D. 2005. Mediator and the mechanism of transcriptional activation. *Trends Biochem Sci*. 30:235-9.
- Kornberg, R.D. 2007. The molecular basis of eucaryotic transcription. *Cell Death Differ*. 14:1989-97.
- Kouzarides, T. 2007. Chromatin modifications and their function. *Cell*. 128:693-705.
- Kundu, T.K., V.B. Palhan, Z. Wang, W. An, P.A. Cole, and R.G. Roeder. 2000. Activator-dependent transcription from chromatin in vitro involving targeted histone acetylation by p300. *Mol Cell*. 6:551-61.
- Kung, A.L., V.I. Rebel, R.T. Bronson, L.E. Ch'ng, C.A. Sieff, D.M. Livingston, and T.P. Yao. 2000. Gene dose-dependent control of hematopoiesis and hematologic tumor suppression by CBP. *Genes Dev*. 14:272-7.
- Kwon, S.H., L. Florens, S.K. Swanson, M.P. Washburn, S.M. Abmayr, and J.L. Workman. 2010. Heterochromatin protein 1 (HP1) connects the FACT histone chaperone complex to the phosphorylated CTD of RNA polymerase II. *Genes Dev*. 24:2133-45.
- Kwon, S.H., and J.L. Workman. 2011. The changing faces of HP1: From heterochromatin formation and gene silencing to euchromatic gene expression: HP1 acts as a positive regulator of transcription. *Bioessays*. 33:280-9.
- Lan, F., R.E. Collins, R. De Cegli, R. Alpatov, J.R. Horton, X. Shi, O. Gozani, X. Cheng, and Y. Shi. 2007. Recognition of unmethylated histone H3 lysine 4 links BHC80 to LSD1-mediated gene repression. *Nature*. 448:718-22.
- Lange, M., B. Kaynak, U.B. Forster, M. Tonjes, J.J. Fischer, C. Grimm, J. Schlesinger, S. Just, I. Dunkel, T. Krueger, S. Mebus, H. Lehrach, R. Lurz, J. Gobom, W. Rottbauer, S. Abdelilah-Seyfried, and S. Sperling. 2008. Regulation of muscle development by DPF3, a novel histone acetylation and methylation reader of the BAF chromatin remodeling complex. *Genes Dev*. 22:2370-84.
- Lau, O.D., T.K. Kundu, R.E. Soccio, S. Ait-Si-Ali, E.M. Khalil, A. Vassilev, A.P. Wolffe, Y. Nakatani, R.G. Roeder, and P.A. Cole. 2000. HATs off: selective synthetic inhibitors of the histone acetyltransferases p300 and PCAF. *Mol Cell*. 5:589-95.
- Lee, J.S., E. Smith, and A. Shilatifard. 2010. The language of histone crosstalk. *Cell*. 142:682-5.
- Lee, K.K., and J.L. Workman. 2007. Histone acetyltransferase complexes: one size doesn't fit all. *Nat Rev Mol Cell Biol*. 8:284-95.
- Lessard, J., J.I. Wu, J.A. Ranish, M. Wan, M.M. Winslow, B.T. Staahl, H. Wu, R. Aebersold, I.A. Graef, and G.R. Crabtree. 2007. An essential switch in

- subunit composition of a chromatin remodeling complex during neural development. *Neuron*. 55:201-15.
- Li, B., M. Carey, and J.L. Workman. 2007a. The role of chromatin during transcription. *Cell*. 128:707-19.
- Li, H., S. Ilin, W. Wang, E.M. Duncan, J. Wysocka, C.D. Allis, and D.J. Patel. 2006. Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. *Nature*. 442:91-5.
- Li, X., Y.K. Lee, J.C. Jeng, Y. Yen, D.C. Schultz, H.M. Shih, and D.K. Ann. 2007b. Role for KAP1 serine 824 phosphorylation and sumoylation/desumoylation switch in regulating KAP1-mediated transcriptional repression. *J Biol Chem*. 282:36177-89.
- Li, Z., L. Chen, N. Kabra, C. Wang, J. Fang, and J. Chen. 2009. Inhibition of SUV39H1 methyltransferase activity by DBC1. *J Biol Chem*. 284:10361-6.
- Linares, L.K., R. Kiernan, R. Triboulet, C. Chable-Bessia, D. Latreille, O. Cuvier, M. Lacroix, L. Le Cam, O. Coux, and M. Benkirane. 2007. Intrinsic ubiquitination activity of PCAF controls the stability of the oncoprotein Hdm2. *Nat Cell Biol*. 9:331-8.
- Liu, X., L. Wang, K. Zhao, P.R. Thompson, Y. Hwang, R. Marmorstein, and P.A. Cole. 2008. The structural basis of protein acetylation by the p300/CBP transcriptional coactivator. *Nature*. 451:846-50.
- Lo, W.S., R.C. Trievel, J.R. Rojas, L. Duggan, J.Y. Hsu, C.D. Allis, R. Marmorstein, and S.L. Berger. 2000. Phosphorylation of serine 10 in histone H3 is functionally linked in vitro and in vivo to Gcn5-mediated acetylation at lysine 14. *Mol Cell*. 5:917-26.
- Lomberk, G., D. Bensi, M.E. Fernandez-Zapico, and R. Urrutia. 2006. Evidence for the existence of an HP1-mediated subcode within the histone code. *Nat Cell Biol*. 8:407-15.
- Luger, K. 2003. Structure and dynamic behavior of nucleosomes. *Curr Opin Genet Dev*. 13:127-35.
- Luijsterburg, M.S., C. Dinant, H. Lans, J. Stap, E. Wiernasz, S. Lagerwerf, D.O. Warmerdam, M. Lindh, M.C. Brink, J.W. Dobrucki, J.A. Aten, M.I. Foustieri, G. Jansen, N.P. Dantuma, W. Vermeulen, L.H. Mullenders, A.B. Houtsmuller, P.J. Verschure, and R. van Driel. 2009. Heterochromatin protein 1 is recruited to various types of DNA damage. *J Cell Biol*. 185:577-86.
- Macdonald, N., J.P. Welburn, M.E. Noble, A. Nguyen, M.B. Yaffe, D. Clynes, J.G. Moggs, G. Orphanides, S. Thomson, J.W. Edmunds, A.L. Clayton, J.A. Endicott, and L.C. Mahadevan. 2005. Molecular basis for the recognition of phosphorylated and phosphoacetylated histone h3 by 14-3-3. *Mol Cell*. 20:199-211.
- Maison, C., and G. Almouzni. 2004. HP1 and the dynamics of heterochromatin maintenance. *Nat Rev Mol Cell Biol*. 5:296-304.
- Manteling, K., B.A. Reddy, V. Swaminathan, A.H. Kishore, N.B. Siddappa, G.V. Kumar, G. Nagashankar, N. Natesh, S. Roy, P.P. Sadhale, U. Ranga, C. Narayana, and T.K. Kundu. 2007. Specific inhibition of p300-HAT alters global gene expression and represses HIV replication. *Chem Biol*. 14:645-57.
- Manzo, F., F.P. Tambaro, A. Mai, and L. Altucci. 2009. Histone acetyltransferase inhibitors and preclinical studies. *Expert Opin Ther Pat*. 19:761-74.

- Marrella, V., P.L. Poliani, C. Sobacchi, F. Grassi, and A. Villa. 2008. Of Omenn and mice. *Trends Immunol.* 29:133-40.
- Martin, C., and Y. Zhang. 2005. The diverse functions of histone lysine methylation. *Nat Rev Mol Cell Biol.* 6:838-49.
- Matsuyama, R., I. Takada, A. Yokoyama, S. Fujiyama-Nakamura, N. Tsuji, H. Kitagawa, R. Fujiki, M. Kim, M. Kouzu-Fujita, T. Yano, and S. Kato. 2010. Double PHD fingers protein DPF2 recognizes acetylated histones and suppresses the function of estrogen-related receptor alpha through histone deacetylase 1. *J Biol Chem.* 285:18166-76.
- Matthews, A.G., A.J. Kuo, S. Ramon-Maiques, S. Han, K.S. Champagne, D. Ivanov, M. Gallardo, D. Carney, P. Cheung, D.N. Ciccone, K.L. Walter, P.J. Utz, Y. Shi, T.G. Kutateladze, W. Yang, O. Gozani, and M.A. Oettinger. 2007. RAG2 PHD finger couples histone H3 lysine 4 trimethylation with V(D)J recombination. *Nature.* 450:1106-10.
- McBryant, S.J., X. Lu, and J.C. Hansen. 2010. Multifunctionality of the linker histones: an emerging role for protein-protein interactions. *Cell Res.* 20:519-28.
- Melcher, M., M. Schmid, L. Aagaard, P. Selenko, G. Laible, and T. Jenuwein. 2000. Structure-function analysis of SUV39H1 reveals a dominant role in heterochromatin organization, chromosome segregation, and mitotic progression. *Mol Cell Biol.* 20:3728-41.
- Metzger, E., A. Imhof, D. Patel, P. Kahl, K. Hoffmeyer, N. Friedrichs, J.M. Muller, H. Greschik, J. Kirfel, S. Ji, N. Kunowska, C. Beisenherz-Huss, T. Gunther, R. Buettner, and R. Schule. 2010. Phosphorylation of histone H3T6 by PKC $\beta$ (I) controls demethylation at histone H3K4. *Nature.* 464:792-6.
- Metzger, E., M. Wissmann, N. Yin, J.M. Muller, R. Schneider, A.H. Peters, T. Gunther, R. Buettner, and R. Schule. 2005. LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. *Nature.* 437:436-9.
- Meulmeester, E., and F. Melchior. 2008. Cell biology: SUMO. *Nature.* 452:709-11.
- Min, J., A. Allali-Hassani, N. Nady, C. Qi, H. Ouyang, Y. Liu, F. MacKenzie, M. Vedadi, and C.H. Arrowsmith. 2007. L3MBTL1 recognition of mono- and dimethylated histones. *Nat Struct Mol Biol.* 14:1229-30.
- Mitchell, J.A., and P. Fraser. 2008. Transcription factories are nuclear subcompartments that remain in the absence of transcription. *Genes Dev.* 22:20-5.
- Mizuguchi, G., X. Shen, J. Landry, W.H. Wu, S. Sen, and C. Wu. 2004. ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. *Science.* 303:343-8.
- Moriniere, J., S. Rousseaux, U. Steuerwald, M. Soler-Lopez, S. Curtet, A.L. Vitte, J. Govin, J. Gaucher, K. Sadoul, D.J. Hart, J. Krijgsvelde, S. Khochbin, C.W. Muller, and C. Petosa. 2009. Cooperative binding of two acetylation marks on a histone tail by a single bromodomain. *Nature.* 461:664-8.
- Mullighan, C.G., J. Zhang, L.H. Kasper, S. Lerach, D. Payne-Turner, L.A. Phillips, S.L. Heatley, L. Holmfeldt, J.R. Collins-Underwood, J. Ma, K.H. Buetow, C.H. Pui, S.D. Baker, P.K. Brindle, and J.R. Downing. 2011. CREBBP mutations in relapsed acute lymphoblastic leukaemia. *Nature.* 471:235-9.

- Nathan, D., K. Ingvarsdottir, D.E. Sterner, G.R. Bylebyl, M. Dokmanovic, J.A. Dorsey, K.A. Whelan, M. Krsmanovic, W.S. Lane, P.B. Meluh, E.S. Johnson, and S.L. Berger. 2006. Histone sumoylation is a negative regulator in *Saccharomyces cerevisiae* and shows dynamic interplay with positive-acting histone modifications. *Genes Dev.* 20:966-76.
- Nelson, C.J., H. Santos-Rosa, and T. Kouzarides. 2006. Proline isomerization of histone H3 regulates lysine methylation and gene expression. *Cell.* 126:905-16.
- Nielsen, P.R., D. Nietlispach, H.R. Mott, J. Callaghan, A. Bannister, T. Kouzarides, A.G. Murzin, N.V. Murzina, and E.D. Laue. 2002. Structure of the HP1 chromodomain bound to histone H3 methylated at lysine 9. *Nature.* 416:103-7.
- Nouspikel, T. 2008. Nucleotide excision repair and neurological diseases. *DNA Repair (Amst).* 7:1155-67.
- O'Connell, S., L. Wang, S. Robert, C.A. Jones, R. Saint, and R.S. Jones. 2001. Polycomblike PHD fingers mediate conserved interaction with enhancer of zeste protein. *J Biol Chem.* 276:43065-73.
- Org, T., F. Chignola, C. Hetenyi, M. Gaetani, A. Rebane, I. Liiv, U. Maran, L. Mollica, M.J. Bottomley, G. Musco, and P. Peterson. 2008. The autoimmune regulator PHD finger binds to non-methylated histone H3K4 to activate gene expression. *EMBO Rep.* 9:370-6.
- Owen, D.J., P. Ornaghi, J.C. Yang, N. Lowe, P.R. Evans, P. Ballario, D. Neuhaus, P. Filetici, and A.A. Travers. 2000. The structural basis for the recognition of acetylated histone H4 by the bromodomain of histone acetyltransferase gcn5p. *EMBO J.* 19:6141-9.
- Paggetti, J., A. Largeot, R. Aucagne, A. Jacquiel, B. Lagrange, X.J. Yang, E. Solary, J.N. Bastie, and L. Delva. 2010. Crosstalk between leukemia-associated proteins MOZ and MLL regulates HOX gene expression in human cord blood CD34+ cells. *Oncogene.* 29:5019-31.
- Park, Y.J., J.V. Chodaparambil, Y. Bao, S.J. McBryant, and K. Luger. 2005. Nucleosome assembly protein 1 exchanges histone H2A-H2B dimers and assists nucleosome sliding. *J Biol Chem.* 280:1817-25.
- Park, Y.J., and K. Luger. 2008. Histone chaperones in nucleosome eviction and histone exchange. *Curr Opin Struct Biol.* 18:282-9.
- Pasqualucci, L., D. Dominguez-Sola, A. Chiarenza, G. Fabbri, A. Grunn, V. Trifonov, L.H. Kasper, S. Lerach, H. Tang, J. Ma, D. Rossi, A. Chadburn, V.V. Murty, C.G. Mullighan, G. Gaidano, R. Rabadan, P.K. Brindle, and R. Dalla-Favera. 2011. Inactivating mutations of acetyltransferase genes in B-cell lymphoma. *Nature.* 471:189-95.
- Patel, J.H., Y. Du, P.G. Ard, C. Phillips, B. Carella, C.J. Chen, C. Rakowski, C. Chatterjee, P.M. Lieberman, W.S. Lane, G.A. Blobel, and S.B. McMahon. 2004. The c-MYC oncoprotein is a substrate of the acetyltransferases hGCN5/PCAF and TIP60. *Mol Cell Biol.* 24:10826-34.
- Pelletier, N., N. Champagne, S. Stifani, and X.J. Yang. 2002. MOZ and MORF histone acetyltransferases interact with the Runt-domain transcription factor Runx2. *Oncogene.* 21:2729-40.

- Perez-Campo, F.M., J. Borrow, V. Kouskoff, and G. Lacaud. 2009. The histone acetyl transferase activity of monocytic leukemia zinc finger is critical for the proliferation of hematopoietic precursors. *Blood*. 113:4866-74.
- Peters, A.H., D. O'Carroll, H. Scherthan, K. Mechtler, S. Sauer, C. Schofer, K. Weipoltshammer, M. Pagani, M. Lachner, A. Kohlmaier, S. Opravil, M. Doyle, M. Sibilia, and T. Jenuwein. 2001. Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. *Cell*. 107:323-37.
- Peterson, C.L., and M.A. Laniel. 2004. Histones and histone modifications. *Curr Biol*. 14:R546-51.
- Petrij, F., R.H. Giles, H.G. Dauwerse, J.J. Saris, R.C. Hennekam, M. Masuno, N. Tommerup, G.J. van Ommen, R.H. Goodman, D.J. Peters, and et al. 1995. Rubinstein-Taybi syndrome caused by mutations in the transcriptional co-activator CBP. *Nature*. 376:348-51.
- Piacentini, L., L. Fanti, M. Berloco, B. Perrini, and S. Pimpinelli. 2003. Heterochromatin protein 1 (HP1) is associated with induced gene expression in *Drosophila* euchromatin. *J Cell Biol*. 161:707-14.
- Piacentini, L., L. Fanti, R. Negri, V. Del Vescovo, A. Fatica, F. Altieri, and S. Pimpinelli. 2009. Heterochromatin protein 1 (HP1a) positively regulates euchromatic gene expression through RNA transcript association and interaction with hnRNPs in *Drosophila*. *PLoS Genet*. 5:e1000670.
- Poux, A.N., and R. Marmorstein. 2003. Molecular basis for Gcn5/PCAF histone acetyltransferase selectivity for histone and nonhistone substrates. *Biochemistry*. 42:14366-74.
- Puri, P.L., V. Sartorelli, X.J. Yang, Y. Hamamori, V.V. Ogryzko, B.H. Howard, L. Kedes, J.Y. Wang, A. Graessmann, Y. Nakatani, and M. Levrero. 1997. Differential roles of p300 and PCAF acetyltransferases in muscle differentiation. *Mol Cell*. 1:35-45.
- Qian, C., and M.M. Zhou. 2006. SET domain protein lysine methyltransferases: Structure, specificity and catalysis. *Cell Mol Life Sci*. 63:2755-63.
- Ramakrishnan, V., J.T. Finch, V. Graziano, P.L. Lee, and R.M. Sweet. 1993. Crystal structure of globular domain of histone H5 and its implications for nucleosome binding. *Nature*. 362:219-23.
- Reed-Inderbitzin, E., I. Moreno-Miralles, S.K. Vanden-Eynden, J. Xie, B. Lutterbach, K.L. Durst-Goodwin, K.S. Luce, B.J. Irvin, M.L. Cleary, S.J. Brandt, and S.W. Hiebert. 2006. RUNX1 associates with histone deacetylases and SUV39H1 to repress transcription. *Oncogene*. 25:5777-86.
- Reinke, H., P.D. Gregory, and W. Horz. 2001. A transient histone hyperacetylation signal marks nucleosomes for remodeling at the PHO8 promoter in vivo. *Mol Cell*. 7:529-38.
- Rokudai, S., Y. Aikawa, Y. Tagata, N. Tsuchida, Y. Taya, and I. Kitabayashi. 2009. Monocytic leukemia zinc finger (MOZ) interacts with p53 to induce p21 expression and cell-cycle arrest. *J Biol Chem*. 284:237-44.
- Routh, A., S. Sandin, and D. Rhodes. 2008. Nucleosome repeat length and linker histone stoichiometry determine chromatin fiber structure. *Proc Natl Acad Sci U S A*. 105:8872-7.

- Ruthenburg, A.J., C.D. Allis, and J. Wysocka. 2007a. Methylation of lysine 4 on histone H3: intricacy of writing and reading a single epigenetic mark. *Mol Cell*. 25:15-30.
- Ruthenburg, A.J., H. Li, T.A. Milne, S. Dewell, R.K. McGinty, M. Yuen, B. Ueberheide, Y. Dou, T.W. Muir, D.J. Patel, and C.D. Allis. 2011. Recognition of a Mononucleosomal Histone Modification Pattern by BPTF via Multivalent Interactions. *Cell*. 145:692-706.
- Ruthenburg, A.J., H. Li, D.J. Patel, and C.D. Allis. 2007b. Multivalent engagement of chromatin modifications by linked binding modules. *Nat Rev Mol Cell Biol*. 8:983-94.
- Ryan, C.M., J.C. Harries, K.B. Kindle, H.M. Collins, and D.M. Heery. 2006. Functional interaction of CREB binding protein (CBP) with nuclear transport proteins and modulation by HDAC inhibitors. *Cell Cycle*. 5:2146-52.
- Saksouk, N., N. Avvakumov, K.S. Champagne, T. Hung, Y. Doyon, C. Cayrou, E. Paquet, M. Ullah, A.J. Landry, V. Cote, X.J. Yang, O. Gozani, T.G. Kutateladze, and J. Cote. 2009. HBO1 HAT complexes target chromatin throughout gene coding regions via multiple PHD finger interactions with histone H3 tail. *Mol Cell*. 33:257-65.
- Sambrook J, R.D. 1989. Molecular cloning a laboratory manual. CSHL press.
- Sanchez, R., and M.M. Zhou. 2011. The PHD finger: a versatile epigenome reader. *Trends Biochem Sci*.
- Santos-Rosa, H., A. Kirmizis, C. Nelson, T. Bartke, N. Saksouk, J. Cote, and T. Kouzarides. 2009. Histone H3 tail clipping regulates gene expression. *Nat Struct Mol Biol*. 16:17-22.
- Santos-Rosa, H., E. Valls, T. Kouzarides, and M. Martinez-Balbas. 2003. Mechanisms of P/CAF auto-acetylation. *Nucleic Acids Res*. 31:4285-92.
- Sapountzi, V., and J. Cote. 2010. MYST-family histone acetyltransferases: beyond chromatin. *Cell Mol Life Sci*.
- Saxowsky, T.T., and P.W. Doetsch. 2006. RNA polymerase encounters with DNA damage: transcription-coupled repair or transcriptional mutagenesis? *Chem Rev*. 106:474-88.
- Schindler, U., H. Beckmann, and A.R. Cashmore. 1993. HAT3.1, a novel Arabidopsis homeodomain protein containing a conserved cysteine-rich region. *Plant J*. 4:137-50.
- Schwarz, K., G.H. Gauss, L. Ludwig, U. Pannicke, Z. Li, D. Lindner, W. Friedrich, R.A. Seger, T.E. Hansen-Hagge, S. Desiderio, M.R. Lieber, and C.R. Bartram. 1996. RAG mutations in human B cell-negative SCID. *Science*. 274:97-9.
- Seet, B.T., I. Dikic, M.M. Zhou, and T. Pawson. 2006. Reading protein modifications with interaction domains. *Nat Rev Mol Cell Biol*. 7:473-83.
- Selvi, R.B., and T.K. Kundu. 2009. Reversible acetylation of chromatin: implication in regulation of gene expression, disease and therapeutics. *Biotechnol J*. 4:375-90.
- Shahbazian, M.D., and M. Grunstein. 2007. Functions of site-specific histone acetylation and deacetylation. *Annu Rev Biochem*. 76:75-100.
- Shaw, P.E. 2007. Peptidyl-prolyl cis/trans isomerases and transcription: is there a twist in the tail? *EMBO Rep*. 8:40-5.

- Shi, X., and O. Gozani. 2005. The fellowships of the ING's. *J Cell Biochem.* 96:1127-36.
- Singh, P.B., and S.D. Georgatos. 2002. HP1: facts, open questions, and speculation. *J Struct Biol.* 140:10-6.
- Smallwood, A., J.C. Black, N. Tanese, S. Pradhan, and M. Carey. 2008. HP1-mediated silencing targets Pol II coactivator complexes. *Nat Struct Mol Biol.* 15:318-20.
- Smith, K.T., and J.L. Workman. 2009. Introducing the acetylome. *Nat Biotechnol.* 27:917-9.
- Somers, J., and T. Owen-Hughes. 2009. Mutations to the histone H3 alpha N region selectively alter the outcome of ATP-dependent nucleosome-remodelling reactions. *Nucleic Acids Res.* 37:2504-13.
- Sowa, M.E., E.J. Bennett, S.P. Gygi, and J.W. Harper. 2009. Defining the human deubiquitinating enzyme interaction landscape. *Cell.* 138:389-403.
- Squatrito, M., C. Gorrini, and B. Amati. 2006. Tip60 in DNA damage response and growth control: many tricks in one HAT. *Trends Cell Biol.* 16:433-42.
- Stassen, M.J., D. Bailey, S. Nelson, V. Chinwalla, and P.J. Harte. 1995. The *Drosophila* trithorax proteins contain a novel variant of the nuclear receptor type DNA binding domain and an ancient conserved motif found in other chromosomal proteins. *Mech Dev.* 52:209-23.
- Strahl, B.D., and C.D. Allis. 2000. The language of covalent histone modifications. *Nature.* 403:41-5.
- Svejstrup, J.Q. 2010. The interface between transcription and mechanisms maintaining genome integrity. *Trends Biochem Sci.* 35:333-8.
- Sykes, S.M., H.S. Mellert, M.A. Holbert, K. Li, R. Marmorstein, W.S. Lane, and S.B. McMahon. 2006. Acetylation of the p53 DNA-binding domain regulates apoptosis induction. *Mol Cell.* 24:841-51.
- Takagi, Y., and R.D. Kornberg. 2006. Mediator as a general transcription factor. *J Biol Chem.* 281:80-9.
- Tanner, K.G., M.R. Langer, and J.M. Denu. 2000a. Kinetic mechanism of human histone acetyltransferase P/CAF. *Biochemistry.* 39:11961-9.
- Tanner, K.G., M.R. Langer, Y. Kim, and J.M. Denu. 2000b. Kinetic mechanism of the histone acetyltransferase GCN5 from yeast. *J Biol Chem.* 275:22048-55.
- Taverna, S.D., H. Li, A.J. Ruthenburg, C.D. Allis, and D.J. Patel. 2007. How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nat Struct Mol Biol.* 14:1025-40.
- Thomas, M.C., C.M. Chiang. 2006. The general transcription machinery and general cofactors. *Crit Rev Biochem Mol Biol.* 41(3):105-78.
- Thomas, T., L.M. Corcoran, R. Gugasyan, M.P. Dixon, T. Brodnicki, S.L. Nutt, D. Metcalf, and A.K. Voss. 2006. Monocytic leukemia zinc finger protein is essential for the development of long-term reconstituting hematopoietic stem cells. *Genes Dev.* 20:1175-86.
- Thompson, P.R., D. Wang, L. Wang, M. Fulco, N. Pediconi, D. Zhang, W. An, Q. Ge, R.G. Roeder, J. Wong, M. Levrero, V. Sartorelli, R.J. Cotter, and P.A. Cole. 2004. Regulation of the p300 HAT domain via a novel activation loop. *Nat Struct Mol Biol.* 11:308-15.



- Thomson, J.P., P.J. Skene, J. Selfridge, T. Clouaire, J. Guy, S. Webb, A.R. Kerr, A. Deaton, R. Andrews, K.D. James, D.J. Turner, R. Illingworth, and A. Bird. 2010. CpG islands influence chromatin structure via the CpG-binding protein Cfp1. *Nature*. 464:1082-6.
- Tjian, R. 1996. The biochemistry of transcription in eukaryotes: a paradigm for multisubunit regulatory complexes. *Philos Trans R Soc Lond B Biol Sci*. 351:491-9.
- Tourkina, E., P. Gooz, J.C. Oates, A. Ludwicka-Bradley, R.M. Silver, and S. Hoffman. 2004. Curcumin-induced apoptosis in scleroderma lung fibroblasts: role of protein kinase cepsilon. *Am J Respir Cell Mol Biol*. 31:28-35.
- Townsley, F.M., B. Thompson, and M. Bienz. 2004. Pygopus residues required for its binding to Legless are critical for transcription and development. *J Biol Chem*. 279:5177-83.
- Trievel, R.C., J.R. Rojas, D.E. Sterner, R.N. Venkataramani, L. Wang, J. Zhou, C.D. Allis, S.L. Berger, and R. Marmorstein. 1999. Crystal structure and mechanism of histone acetylation of the yeast GCN5 transcriptional coactivator. *Proc Natl Acad Sci U S A*. 96:8931-6.
- Tsai, W.W., Z. Wang, T.T. Yiu, K.C. Akdemir, W. Xia, S. Winter, C.Y. Tsai, X. Shi, D. Schwarzer, W. Plunkett, B. Aronow, O. Gozani, W. Fischle, M.C. Hung, D.J. Patel, and M.C. Barton. 2010. TRIM24 links a non-canonical histone signature to breast cancer. *Nature*. 468:927-32.
- Tschiersch, B., A. Hofmann, V. Krauss, R. Dorn, G. Korge, and G. Reuter. 1994. The protein encoded by the *Drosophila* position-effect variegation suppressor gene Su(var)3-9 combines domains of antagonistic regulators of homeotic gene complexes. *EMBO J*. 13:3822-31.
- Ullah, M., N. Pelletier, L. Xiao, S.P. Zhao, K. Wang, C. Degerny, S. Tahmasebi, C. Cayrou, Y. Doyon, S.L. Goh, N. Champagne, J. Cote, and X.J. Yang. 2008. Molecular architecture of quartet MOZ/MORF histone acetyltransferase complexes. *Mol Cell Biol*. 28:6828-43.
- Umehara, T., Y. Nakamura, M. Wakamori, K. Ozato, S. Yokoyama, and B. Padmanabhan. 2010. Structural implications for K5/K12-di-acetylated histone H4 recognition by the second bromodomain of BRD2. *FEBS Lett*. 584:3901-8.
- van Oostveen, J., J. Bijl, F. Raaphorst, J. Walboomers, and C. Meijer. 1999. The role of homeobox genes in normal hematopoiesis and hematological malignancies. *Leukemia*. 13:1675-90.
- Vaquero, A., M. Scher, H. Erdjument-Bromage, P. Tempst, L. Serrano, and D. Reinberg. 2007. SIRT1 regulates the histone methyl-transferase SUV39H1 during heterochromatin formation. *Nature*. 450:440-4.
- Vaquero, A., M. Scher, D. Lee, H. Erdjument-Bromage, P. Tempst, and D. Reinberg. 2004. Human SirT1 interacts with histone H1 and promotes formation of facultative heterochromatin. *Mol Cell*. 16:93-105.
- Verdone, L., E. Agricola, M. Caserta, and E. Di Mauro. 2006. Histone acetylation in gene regulation. *Brief Funct Genomic Proteomic*. 5:209-21.
- Voss, A.K., C. Collin, M.P. Dixon, and T. Thomas. 2009. Moz and retinoic acid coordinately regulate H3K9 acetylation, Hox gene expression, and segment identity. *Dev Cell*. 17:674-86.

- Wang, G.G., J. Song, Z. Wang, H.L. Dormann, F. Casadio, H. Li, J.L. Luo, D.J. Patel, and C.D. Allis. 2009. Haematopoietic malignancies caused by dysregulation of a chromatin-binding PHD finger. *Nature*. 459:847-51.
- Wierzbicki, A.T., T.S. Ream, J.R. Haag, and C.S. Pikaard. 2009. RNA polymerase V transcription guides ARGONAUTE4 to chromatin. *Nat Genet*. 41:630-4.
- Willis, I.M. 1993. RNA polymerase III. Genes, factors and transcriptional specificity. *Eur J Biochem*. 212:1-11.
- Woodcock, C.L., and R.P. Ghosh. 2010. Chromatin higher-order structure and dynamics. *Cold Spring Harb Perspect Biol*. 2:a000596.
- Xu, M., and P.R. Cook. 2008. The role of specialized transcription factories in chromosome pairing. *Biochim Biophys Acta*. 1783:2155-60.
- Xu, W., D.G. Edmondson, and S.Y. Roth. 1998. Mammalian GCN5 and P/CAF acetyltransferases have homologous amino-terminal domains important for recognition of nucleosomal substrates. *Mol Cell Biol*. 18:5659-69.
- Yamaguchi, Y., T. Takagi, T. Wada, K. Yano, A. Furuya, S. Sugimoto, J. Hasegawa, and H. Handa. 1999. NELF, a multisubunit complex containing RD, cooperates with DSIF to repress RNA polymerase II elongation. *Cell*. 97:41-51.
- Yan, Y., S. Harper, D.W. Speicher, and R. Marmorstein. 2002. The catalytic mechanism of the ESA1 histone acetyltransferase involves a self-acetylated intermediate. *Nat Struct Biol*. 9:862-9.
- Yang, X.J. 2004. The diverse superfamily of lysine acetyltransferases and their roles in leukemia and other diseases. *Nucleic Acids Res*. 32:959-76.
- Yang, X.J., and M. Ullah. 2007. MOZ and MORF, two large MYSTic HATs in normal and cancer stem cells. *Oncogene*. 26:5408-19.
- Yoshida, H., and I. Kitabayashi. 2008. Chromatin regulation by AML1 complex. *Int J Hematol*. 87:19-24.
- Yu, A., H.Y. Fan, D. Liao, A.D. Bailey, and A.M. Weiner. 2000. Activation of p53 or loss of the Cockayne syndrome group B repair protein causes metaphase fragility of human U1, U2, and 5S genes. *Mol Cell*. 5:801-10.
- Yuan, L.W., and A. Giordano. 2002. Acetyltransferase machinery conserved in p300/CBP-family proteins. *Oncogene*. 21:2253-60.
- Zeng, L., K.L. Yap, A.V. Ivanov, X. Wang, S. Mujtaba, O. Plotnikova, F.J. Rauscher, 3rd, and M.M. Zhou. 2008a. Structural insights into human KAP1 PHD finger-bromodomain and its role in gene silencing. *Nat Struct Mol Biol*. 15:626-33.
- Zeng, L., Q. Zhang, G. Gerona-Navarro, N. Moshkina, and M.M. Zhou. 2008b. Structural basis of site-specific histone recognition by the bromodomains of human coactivators PCAF and CBP/p300. *Structure*. 16:643-52.
- Zeng, L., Q. Zhang, S. Li, A.N. Plotnikov, M.J. Walsh, and M.M. Zhou. 2010. Mechanism and regulation of acetylated histone binding by the tandem PHD finger of DPF3b. *Nature*. 466:258-62.
- Zheng, Y., K. Balasubramanyam, M. Cebrat, D. Buck, F. Guidez, A. Zelent, R.M. Alani, and P.A. Cole. 2005. Synthesis and evaluation of a potent and selective cell-permeable p300 histone acetyltransferase inhibitor. *J Am Chem Soc*. 127:17182-3.

- Zheng, Y.G., J. Wu, Z. Chen, and M. Goodman. 2008. Chemical regulation of epigenetic modifications: opportunities for new cancer therapy. *Med Res Rev.* 28:645-87.
- Zhovmer, A., V. Oksenych, and F. Coin. 2010. Two sides of the same coin: TFIIH complexes in transcription and DNA repair. *ScientificWorldJournal.* 10:633-43.